

二、研究計畫內容

(一) 摘要

書法，一種漢字的書寫方法，也是一門中國藝術。依照文字特點與其中文字含義，以筆法、文字結構與其章法書寫。在文字的結構規律之下變化，成為富有美感的藝術作品。另外，書法具有悠久、深厚的傳統文化，是琴棋書畫四藝之一。中國歷代，產生了許多大名鼎鼎的書法家，例如王羲之、顏真卿、柳公權等等，也留下了許多代表作品。現代，台灣書法教育越來越少，越來越少人接觸書法藝術，反觀鄰近的日本、韓國與中國已經開始推廣這項傳統文化並且重視這個議題，中國在近年已全面推廣書法教育，日本「書道」，韓國「書藝」。本計畫試圖從書法文字著手，期望能夠提出基於深度學習架構中文字的筆順系統。深度學習是一種以人工神經網路為架構，已被應用在電腦視覺、影像處理各個領域上有重大的貢獻。本計畫中期望運用此技術結合相關多媒體技術，利用 Local Model 完成單一筆畫的起始點與結束點，到達筆畫的結束點時，再使用 Global Model 預測下一筆畫的起始點，以此架構發展辨識書法字筆畫系統。

關鍵字：書法字、深度學習、筆順學習

(二) 研究動機與研究問題

書法為我國流傳已久的國粹，根據文化部文化產業提要分析[1] 發現，民眾參與繪畫書法相關活動有逐年遞減的比率，在書法教育方面根據台灣書法基礎教育現況之調查研究[2] 指出，國小書法教育因語文時數縮減，教師缺乏足夠的上課時間進行書法教學。多數教師認為書法教材已不適用，導致教學成效有限。本計畫結合傳統書法藝術與最新的機器學習技術，希望能夠透過科技的方式能夠推廣這項傳統藝術。

文化部 2016-2017 文化統計提要分析[1]

	2013	2014-2015	2015-2016	2016-2017
繪畫書法類	25.3%	25.8%	24.7%	21.7%

繪畫書法類（如水墨、書法與篆刻、水彩、版畫、插畫及素描 等）

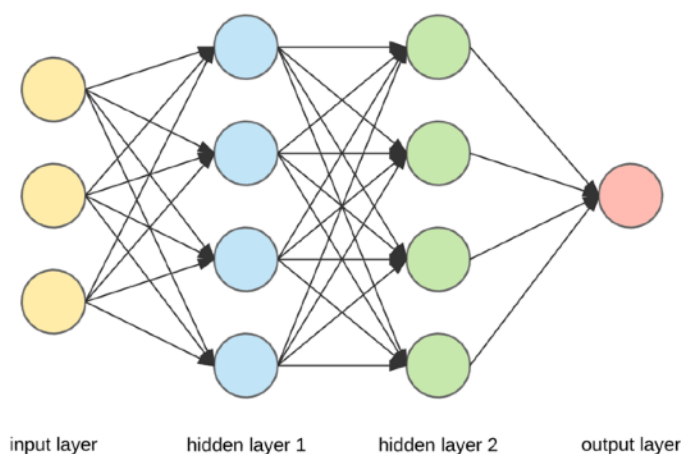
現階段機械手臂無法直接對二維圖像進行書寫，必須要有序列的資料，才可以讓機械手臂書寫，希望透過目前熱門的深度學習技術，可以從二維圖像中，產生出筆順資料的序列，透過得到的筆順資料，讓機械手臂可以正確書寫出結果。



圖一、機械手臂寫書法圖

目前研究中文書法筆畫的研究的相關研究很少，而且成效不顯著，Drawing Order Recovery based on deep learning[3] 和 Reconstructing stroke and writing sequence from Chinese character image[4] 皆是重新建構出中文筆畫的相關研究，但是兩者的成效在10畫以上的筆畫成功率相當的低，對於中文字來說筆畫常常超過10劃以上，增加此計畫的困難度。

深度學習是近年來熱門的而且具有相當發展潛力的技術，在電腦視覺與影像處理相關的研究領域，有許多論文開始應用深度學習在其中，解決複雜的問題。深度學習是基於神經網路架構，用來模擬人類大腦的思考方式，從低階影像像素表示，到中階邊緣與形狀等等概念，最後進而達到高階學習的意涵，能夠辨識物件、人臉等等，如果遇到未看過的全新影像，深度學習的突破能夠產生相對應的圖像敘述，對影像「瞭解」而做出辨識上的判斷。



圖二、深度學習架構圖

此研究探討的問題主要有兩點：

- 中文字複雜結構

中文字在結構方面相當的複雜且多樣，以「永」字為例，就可以分成八種不同的筆畫（側、勒、弩（又作努）、趯、策、掠、啄、磔），且每種筆畫書寫的方式都大不相同，欲從中文字拆解出每個筆畫的分類，以及正確寫出筆畫，都是必須克服的問題。

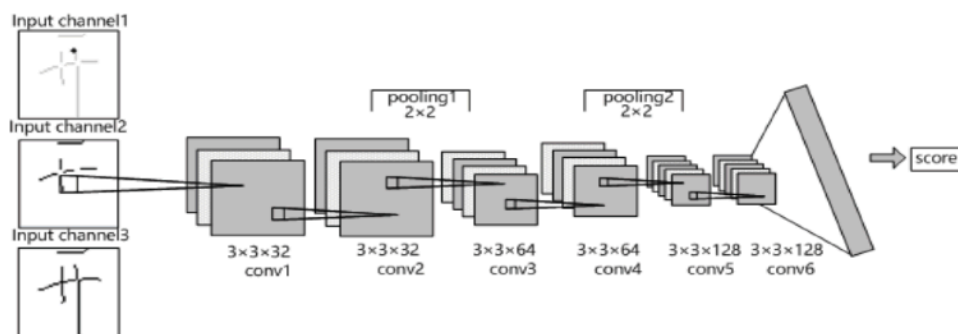
- 中文字筆畫差異巨大

從過去的研究結果得知，對於筆畫較少的字，取得高準確度的筆順較容易，但是中文筆畫數具有相當大的落差，從一劃到四十劃都有可能，當文字的筆畫數增加，能正確寫出字的筆順的難度也隨之增加。

（三） 文獻回顧與探討

目前關於字體重構大多是採用CNN作為訓練網路，以Drawing order recovery based on deep learning[10] 這篇論文為例，將整個過程分為兩個部分，分別是筆畫重構的網路訓練以及中文字筆順重構。第一部分將訓練集圖像轉化為48*48 的大小，對其進行細線化，保留字體骨架。以三個通道（分別為，當前書寫狀態，已完成的書寫狀態與完整字體），透過卷積神經網路 (Convolutional Neural Networks)，可以預測下筆畫的位置，直到整個字體筆順完成，但是以圖二的實驗結果發現，筆畫在10 畫以上的中文字辨識準確度並不高。

在筆順學習方面，我們參考 Teaching robot to draw [6] 這篇論文，而論文中提到，傳統的 motion planning methods 對於機器手臂寫字是不足夠的，以 Rapidly-Exploring Random Trees: A New Tool for Path Planning[10] 為例，以初始點當作根節點，通過隨機的取樣的方式增加子節點，可以得到一個隨機拓展樹，當樹中的子節點包含了目標點或進入了目標區域，便可以在樹中找到一條由從初始點到目標點的路徑，此路徑即是用於機器手臂書寫的路徑，但是所找到的路徑可能不會考慮到每個筆畫連續性。



圖三、Drawing order recovery based on deep learning CNN 架構[10]

因此論文的作者提出，每個字是由多個 local part 所組成，global 則是表示了整個字的所有 part，如此，在每個 local part 裡面的移動，都不會造成 global reconfiguration，所以就能將所有的 drawing motion，以 local 的移動來表示，每個 local part 都有各自對應的 local movement，而機器手臂有了這些資料，就可以正確地將每個筆畫完成。

而目前現有的機器書寫的方法大多集中在圖像處理的技術，用來預測圖像上的移動，但是這種方法缺少學習筆順規則。Robotic calligraphy - Learning how to write single strokes of Chinese and Japanese characters [11] 透過文字的形狀將其區分為不同的筆畫，但是此做法的成果好壞，很大程度取決於手動提供的輸入資料的位置，而且在書法中有些筆畫的起始點與結束點之間有複雜的動作，雖然最後的結果相近，但是在書寫過程筆順軌跡有很大的不同。

Artist agent: A reinforcement learning approach to automatic stroke generation in oriental ink painting [12] 使用了強化學習的方法，將問題轉為 Markov Decision Process (MDP)，相較於以前的方法，需要更少監督，只是在訓練過程中，需要大量的數據，還需要手動標記每個筆畫的起始點與結束點，相當的耗費成本。

(四) 研究方法及步驟

本計畫預定執行過程分為兩個部分，分別為筆畫學習網路訓練，以及筆畫學習。筆畫學習網路訓練部分，我們採用 KanjiVG[8] 資料集，為了讓程式能夠有高精準度的預測筆畫順序，採用影像資料處理與神經網路學習，建構出判斷正確的筆畫模型。筆畫學習使用手機攝像鏡頭或者是上傳書法字圖檔，將其細線化以後，經由訓練好的神經網路模型對檔案進行筆順預測。

神經網路學習

當影像資料處理完後，這個部分我們採用 Teaching robot to draw[6]，整個網路架構分為 Local Model 與 Global Model 兩個部分，Local Model，主要的目的是從單一筆畫起始點開始，能夠找到結束點完成單一筆畫，此階段使用 ResNet 與 LSTM 進行預測，當筆順到達端點時，使

用Global Model，從當前筆畫的結束點預測下一筆畫的起始點，這個部分使用ResNet 預測下一筆畫的起始點。

Local Model：首先將文字圖片分為四個通道：

1. 目前位置
2. 已經拜訪過的區域： X_t^{Lenv} (圖四b)
3. 與目前位置所連接的區域： X_t^{Lcon} (圖四c)
4. 未拜訪的區域： X_t^{Ldif} (圖四d)

S_t^L 為已經拜訪過的區域、與目前位置所連接的區域、未拜訪過的區域三者的集合。除了目前位置之外，將另外三個通道作為輸入進入Residual Network，Residual Network 由四個residual blocks 組成，每個residual block 下還有兩個sub-blocks，sub-block 由batch normalization layer、rectified linear unit 和 two-dimensional convolutional layer組成。two-dimensional convolutional layer的通道為[16,16],[16,16],[16,32],[32,64]。經由ResNet後，會形成(100,100,64)的張量。

$$e_t^L = f_{\theta_{Residual}}(S_t^L)$$

再提取出以目前位置為中心，5x5 範圍內的資訊，將其拓展為 $5 \times 5 \times 64 = 1600$ 的向量進入LSTM。

$$V_t^L = \text{reshape}(e_t^{L*}), \quad c_t^L, h_t^L = f_{\theta_{LSTM}}([v_t^L; h_{t-1}^L])$$

利用LSTM預測是否要繼續當前的筆順，或者是已經到達端點，如果到達端點，進入Global Model，預測下一筆畫的起始位置。LSTM主要的目的是預測筆順下一個像素點，以確保預測連續性，舉例來說，如果像素點是從左邊過來，藉由LSTM預測接下來應該要往右邊走，不會破壞筆畫連續性。

Global Model：如果 Local Model touch action 為0 時進入Global Mode，目的是要預測下一筆畫得開頭位置，ResNet 所採用的權重參數與Local Model 相同。將四個通道：

1. 目前位置： X_t^{Gloc}
2. 已經拜訪過的區域： X_t^{Genv} (圖四f)
3. 與目前位置所連接的區域： X_t^{Glast} (圖四g)
4. 未拜訪的區域： X_t^{Gdif} (圖四h)

S_t^G 為已經拜訪過的區域、與目前位置所連接的區域、未拜訪過的區域四者的集合。將上述的集合輸入 ResNet，原本 (100, 100, 4) 的資料變成 (100, 100, 64) 的張量

$$e_t^G = f_{\theta_{Residual}}(S_t^G)$$

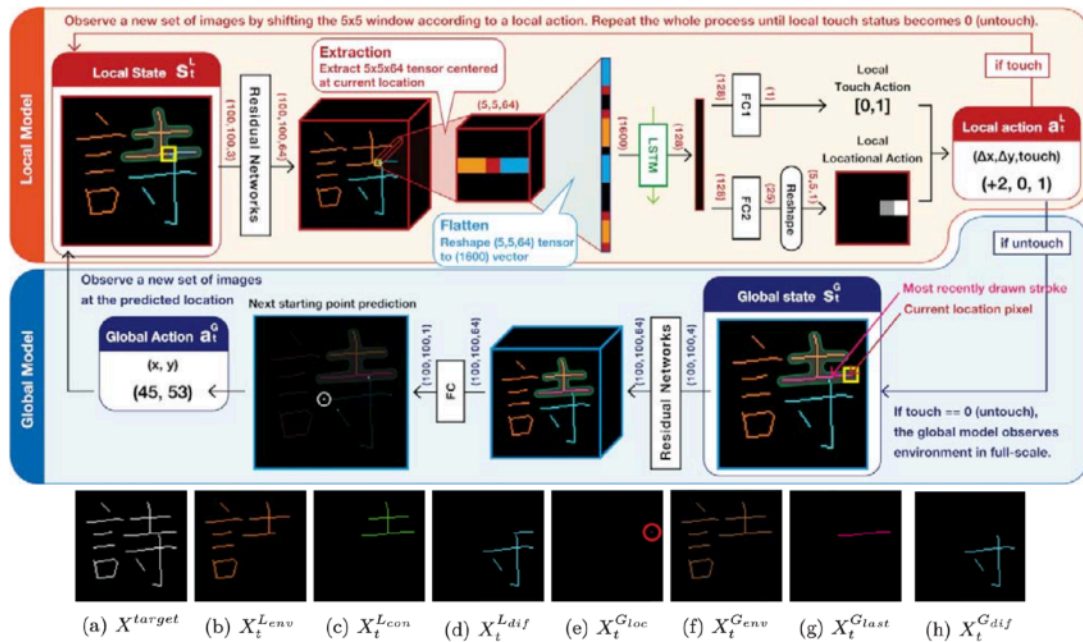
將經過 ResNet 的結果，當成 fully-connected layer 的輸入，可以得到 (100, 100, 1) 的輸出。

$$c_t^G = f_{\theta_{FC}}(e_t^G)$$

利用fully-connected layer 的輸出結果，可以用來得到下一筆畫的起始位置。

$$a_t^G = \arg \max_{x(y)} c_t^G$$

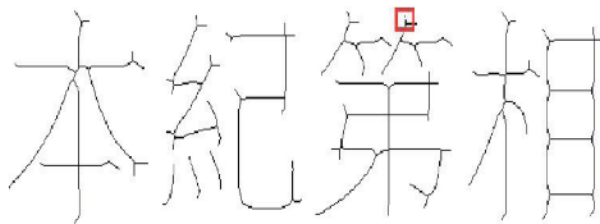
再將起始位置傳入 Local model，完成下一筆畫，而重複如此的過程，就能從原本的二維文字圖檔，得到文字的序列資料，機械手臂也能透過此資料，完整書寫出圖檔中的文字，並且包含了文字的筆畫順序。



圖四、Teaching Robot To Draw[6]

影像資料處理

在進行筆順提取之前，我們使用細線化影像處理，對字提取出骨架，我們採用An Improved Image Thinning Algorithm and Its Application in Chinese Character Image Refining[7]，這個做法是基於OPTA thinning Algorithm的優化，OPTA thinning Algorithm取出目標點周圍8個pixel並且定義出兩種template，分別為delete templates和retention templates，如果目標點符和兩種template則保留反之則刪除，如果不符合delete template則保留此點。但是，OPTA thinning algorithm所得到的文字結果會有毛邊以及交接點有變形的問題，而An Improved Image Thinning Algorithm and Its Application in Chinese Character [7]此篇論文所提出的方法中，增加取的目標點周圍pixel，以及修改與增加 template，可以減少毛邊以及交接點變形的問題，還原文字的真實度，可以用於解決上述的問題。

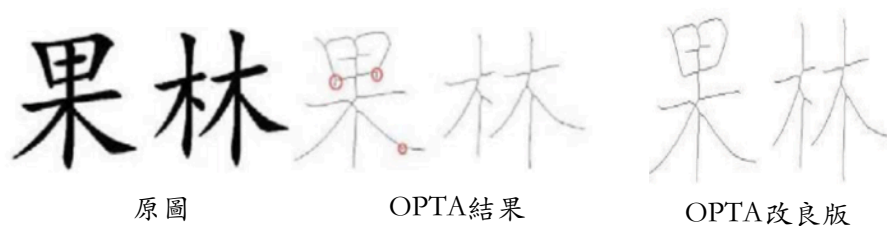


圖五、OPTA文字結果

細線化處理

首先，定義出對目標點右邊和下方像素點的Equivalent non-deletable template 與simple connected invariant template，為了要減少毛邊，所以定義了Burr removal template，來減少毛邊的問題，最後為了需要保留一些小區塊定義Small area retention template。

將先前採用的資料庫當作輸入資料，將圖像每個像素與其8鄰點定義為non-deletable point and deletable point，先後針對前面所定義的Equivalent non-deletable template 與simple connected invariant template做比對，如果該像素的右方與下方像素點皆為 non-deletable point 或者是simple connected invariant template，則刪除該點。接下來比對Burr removal template與



圖六、OPTA改良後差異

資料集

使用KanjiVG[8] 資料集，共有11458個包含中文、日本漢字、英文字組成的資料集，每筆資料具有文字圖檔與其對應筆畫順序檔案。



圖七、KanjiVG[8] 資料集

研究方法步驟

1. 讀取文字影像圖檔
2. 使用An Improved Image Thinning Algorithm and Its Application in Chinese Character Image Refining[7]細線化處理文字圖像，減少毛邊以及變形。
3. 將細線化後的資料送進Teaching robot to draw[6]深度學習模型，從Local Model 預測筆順連續性，Global Model 預測筆畫順序起點。
4. 將資料回饋給建立的模型，用以提升系統準確率

(五) 預期結果

以 Teaching robot to draw [6] 此篇論文的成功得知，當文字的筆畫超過十劃時，成功率約只有十六個百分比，我們希望能夠提出改良的方法，讓機器手臂書寫文字的成功率能有所提升。

根據計畫內容，在短期目標我們希望能提升筆順學習的成功率。中期目標希望能夠普及字體，不受字體、寫字風格的限制。長期的目標主要是想要讓這個系統普及化，運用在學習方面，例如中小學的書法課程上，讓這方面的國粹在未來更多人認識。

實現這個系統主要是能夠擷取書法字，經過設計好的演算法辨識出正確的筆順，再把正確的筆順結果產出，透過這樣的設計能夠幫助學習書法者更了解書法字書寫。

	Number of strokes		
	0~5	5~10	>10
Quantity	289	607	104
$D_n U_n$	67.8%	53.6%	12.8%
$D_n U_y$	78.3%	60.9%	18.85
Average	73.0%	57.2%	15.8%

圖八、預測準確度[9]

(六) 參考文獻

1. 李秀華, 楊勢年, 台灣書法基礎教育現況之調查研究, 2008
2. 文化部- 2016-2017文化產業提要分析
3. Bocheng Zhao, Minghao Yang, Jianhua Tao, Drawing Order Recovery based on deep learning, 2019 Eleventh International Conference on Advanced Computational Intelligence (ICACI), 2019
4. Kai-Tai Tang, Howard Leung, Reconstructing stroke and writing sequence from Chinese character image, 2007 International Conference on Machine Learning and Cybernetics, 2007
5. Feng Lin, Xiaoou Tang, Off-line handwritten Chinese character stroke extraction, Object recognition supported by user interaction for service robots, 2002
6. Atsunobu Kotani and Stefanie Tellex, Teaching Robots To Draw, 2019 International Conference on Robotics and Automation (ICRA), 2019
7. Wang Miaomiao, Li Zhenglin, Si Fuyuan, Guan Lei, An Improved Image Thinning Algorithm and Its Application in Chinese Character Image Refining, 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), 2019
8. Ulrich Apel, VG Kanji, August 2009, [online] Available: <http://kanjivg.tagaini.net/>.
9. Steven M LaValle, Rapidly-exploring random trees: A new tool for path planning, 1998
10. Samuel Mueller, Nico Huebel, Markus Waibel, D'Andrea. Raffaello, Robotic calligraphy-learning how to write single strokes of chinese and japanese characters, 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2013
11. Ning Xie, Hirotaka Hachiya, Masashi Sugiyama, Artist agent: A reinforcement learning approach to automatic stroke generation in oriental ink painting, 2013

(七) 需要指導教授指導內容

本計畫重構文字筆順的問題上，希望可以透過老師的專業知識教導我們，關於影像處理方面的相關演算法與技術，結合深度學習來達成相關領域的指導，尤其是在深度學習方面，並無接觸過這項領域的技術以及實作，對於這部分並不是很熟悉。也希望能夠累積相關資料處理的經驗，透過老師的帶領與指導，一定可以給予我們極大的幫助，達到最好的效果。