

利用機器學習推斷手寫字體風格

系所／資訊工程學系

指導老師／徐熊健

組員／林子恒、黃彥愷、鍾榮輯、張克承

本研究計畫的目的是建立一套機器學習的模型架構系統，藉由輸入足量的字集檔案訓練，使得機器能夠提取足夠的字集特徵並能夠推斷風格。甚至能夠學習人類的手寫字跡，推斷出正確且相似於人類手寫風格的字元出來。

首先將 56 種不同的字體（包含手寫簡字集、手寫簡繁混合字集）以及使用者手寫字集，利用 tesseract 與 jTessBoxEditor 的 OCR 工具包提取使用者的手寫字集做為神經網路的訓練資料，如圖 1 所示。我們首先對於這 56 種字體中各隨機取樣出 1000 個字元，再將取樣出的字元圖片檔與源字體的字元圖片檔配對成單一圖片檔，並將其處理成 obj 檔作為訓練資料以及驗證資料使用，再利用 GAN 神經網路進行批次訓練。

GAN 神經網路訓練分作兩部分，其一為承襲 pix2pix 的理論架構延伸的 zi2zi，訓練字體與字體之間的映射關係，使機器懂得各個筆畫及部首間的映射，進而造出樣貌完整的字元。其二為利用 embedding 層，將每個字元（包

括手寫字）給予特徵空間中隨機的多維度座標，並於訓練過程中將相同字體的字元聚類在一起。即為，在特徵空間中的相同字體的字元彼此的距離會趨於接近。如此一來，模型不但學習了多種字體，並且也學習了分類，增加了字體的多元性及變化性。

最後的微調過程中，我們凍結大部分先前訓練完成的模型參數，只微調全連接層，將模型導向指定的目標字體，換句話說，藉由 56 種字體學習如何「寫字」，再透過使用者字集的特徵，學習字元的「風格」，兩者相輔相成。經由其他字集的「協助」且利用少許的字元，便能推斷出目標字體的全部字元及風格。

圖 2 為字集推斷結果，圖 3 為手寫生成比對（左為使用者手寫字元，右為推斷字元），圖 4 為流程圖。

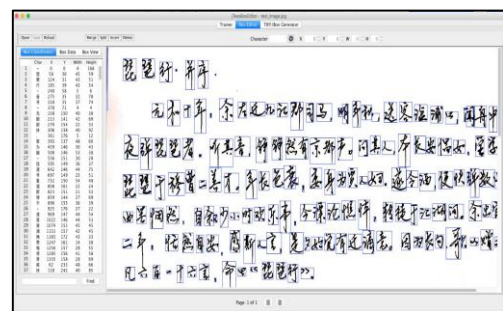


圖 1：使用 OCR 工具包提取手寫字元

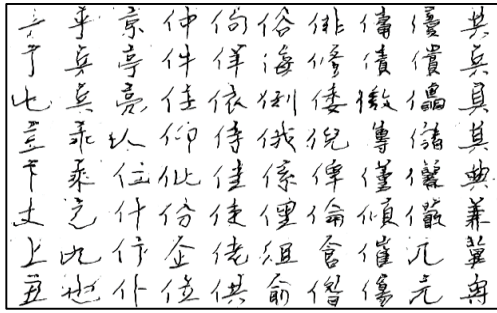


圖 2：字集推斷結果



圖 3-1：手寫生成比對—獨



圖 3-2：手寫生成比對—散

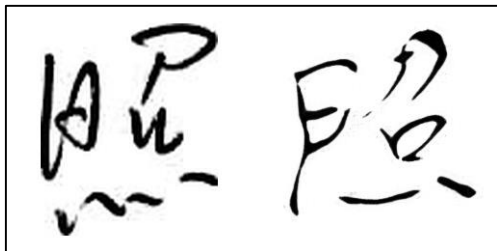


圖 3-3：手寫生成比對—照



圖 3-4：手寫生成比對—霹



圖 3-5：手寫生成比對—侯



圖 3-6：手寫生成比對—鐵



圖 3-7：手寫生成比對—歎

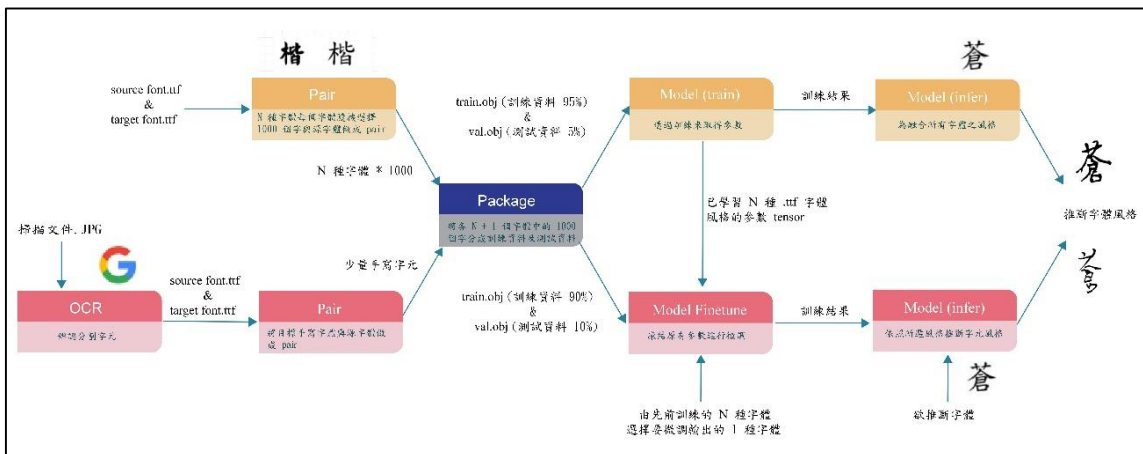


圖 4：流程圖