

# 台科大 109 學年度「彩色影像處理」作業三：簡化的 CNN 數值辨識

孫沛立老師

附件資料:

有 train1000 與 test1000 兩組各 1000 個 0~9 的數值影像  
1~100 號：數值'0'  
101~200 號：數值'1'  
....  
901~1000 號：數值'9'

作業內容：

1. 將 train1000 裡的要訓練的影像縮小至 8x8。
2. 第一層卷積：用兩種 3x3 濾鏡做卷積運算，運算後的特徵圖大小為 8x8x2。可參考 how\_CNNs\_work.ppt 講義。Matlab 可用 `imfilter()` 執行卷積運算。

$$filter1 = \begin{bmatrix} -1 & -1 & 1 \\ -1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix} \quad filter2 \text{ 自行設計}$$

3. 第一層池化：對步驟 2 的結果做 2x2 的最大池化(max pooling)，運算後的特徵圖大小為 4x4x2。  
如果使用 Matlab，可先自己寫一個將 2x2 矩陣的最大值算出的 `max2()` 函式，再用 `blkproc()` 函式，執行分區的 `@max2` 運算。
4. 第二層卷積：用步驟 2 的兩種 3x3 濾鏡做卷積運算，運算後的特徵圖大小為 4x4x2。
5. 第二層池化：對步驟 4 的結果做 2x2 的最大池化(max pooling)，運算後的特徵圖大小為 2x2x2。
6. 平坦化：將 2x2x2 特徵轉成 1x8 的特徵資料，再加常數項 1，成為 1x9 的資料。

相當於  $y = a_1x_1 + a_2x_2 + \dots + a_8x_8 + a_9$ ，其中  $y$  為目標值(0 或 1)， $x$  是特徵值， $a$  是待優化的係數。以下將透過線性迴歸，獲得係數矩陣  $A = [a_1 \ a_2 \ \dots \ a_9]'$ 。

7. 線性迴歸(取代深度學習)：如果用來辨識'0'與'1'，就讀取前 200 幅影像的資料，線性迴歸自變數  $X$  的大小是 200x9，依變數  $Y$  的大小是 200x1， $Y$  的前 100 個數值設為 0，後 100 的數值設為 1。使用最小平方方法  $A = (X^T X)^{-1} (X^T Y)$  算出 9x1 的係數矩陣。若使用 Matlab， $A = \text{inv}(X' * X) * (X' * Y) = X \backslash Y$ 。

8. 測試(混淆矩陣)：

可套用 train 組的 200 個影像特徵資料，或另外用 test 組的 200 個影像特徵資料做評估。首先用  $Y_p = X * A$  獲得預測值  $Y_p$ ，如果前 100 個影像數值低於 0.5，代表 '0' 被正確判斷，請把數量填入混淆矩陣(1,1)的位置。如果前 100 個影像數值高於 0.5，代表 '0' 被辨識為 '1'，請把數量填入混淆矩陣(2,1)的位置。如果後 100 個影像數值高於 0.5，代表 '1' 被正確判斷，請把數量填入混淆矩陣(2,2)的位置。如果後 100 個影像數值低於 0.5，代表 '0' 被辨識為 '1'，請把數量填入混淆矩陣(1,2)的位置。

表 1: 混淆矩陣，以 '0' 與 '1' 數值影像辨識為例

混淆矩陣		實際的類別	
		'0'	'1'
預測的類別	'0'	74	18
	'1'	26	82

9. 隨機測試：從這 200 張影像任意選取(rand())15 張，評估其效果，每個圖的標題含預測的數值，以及隨機選取的影像編號(括弧內)。若答對，文字是藍色的；若答錯，文字是紅色的。如下列例子：

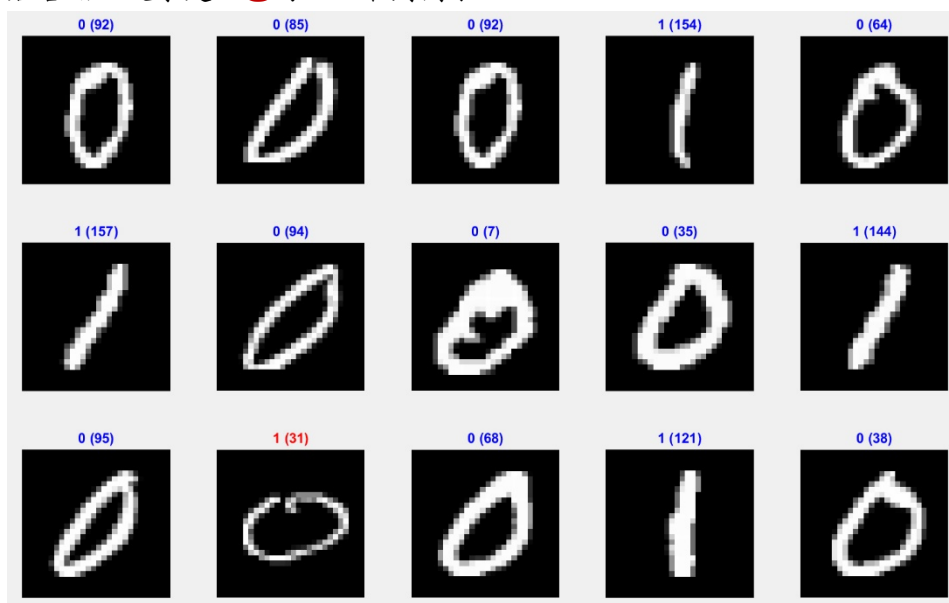


圖 1: 從 200 張測試影像隨機選 15 張顯示辨識成果，以 '0', '1' 為例。

10. 加分題：改寫成能夠讀取任意兩個數值，顯示混淆矩陣，與 15 個隨機選取影像的測試結果。例如下圖：

>> 1st number (0~9): 4

>> 2nd number (0~9): 7

表 2: 混淆矩陣，以'4'與'7'數值影像辨識為例

混淆矩陣		實際的類別	
		'4'	'7'
預測的類別	'4'	89	18
	'7'	11	82

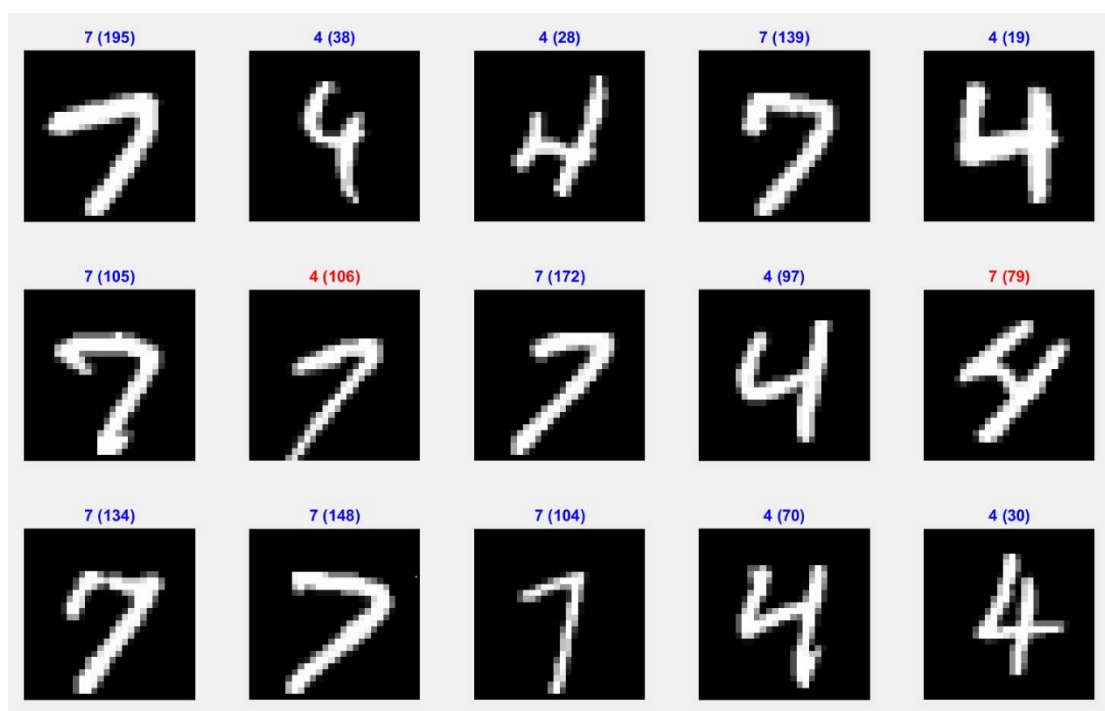


圖 2: 從 200 張測試影像隨機選 15 張顯示辨識成果，以'4','7'為例，紅色為辨識錯誤的影像。

程式語言：可用 Matlab, Python, C++, Java, VB。

繳交內容：附上程式碼，程式需詳細註解，不用附影像檔。程式貼入 Word 檔，連同執行結果製作一個以 HW3\_學號命名的 PDF 檔。將所有檔案放在以 HW3\_學號命名的檔案匣，打包成 zip 檔上傳。

注意：不得使用深度學習或類神經網路套件。

繳交期限：110 年 1 月 15 日 24:00 前上傳至 Moodle 作業區。

有問題可請教色彩所江晉豪同學(ewan21621@gmail.com)