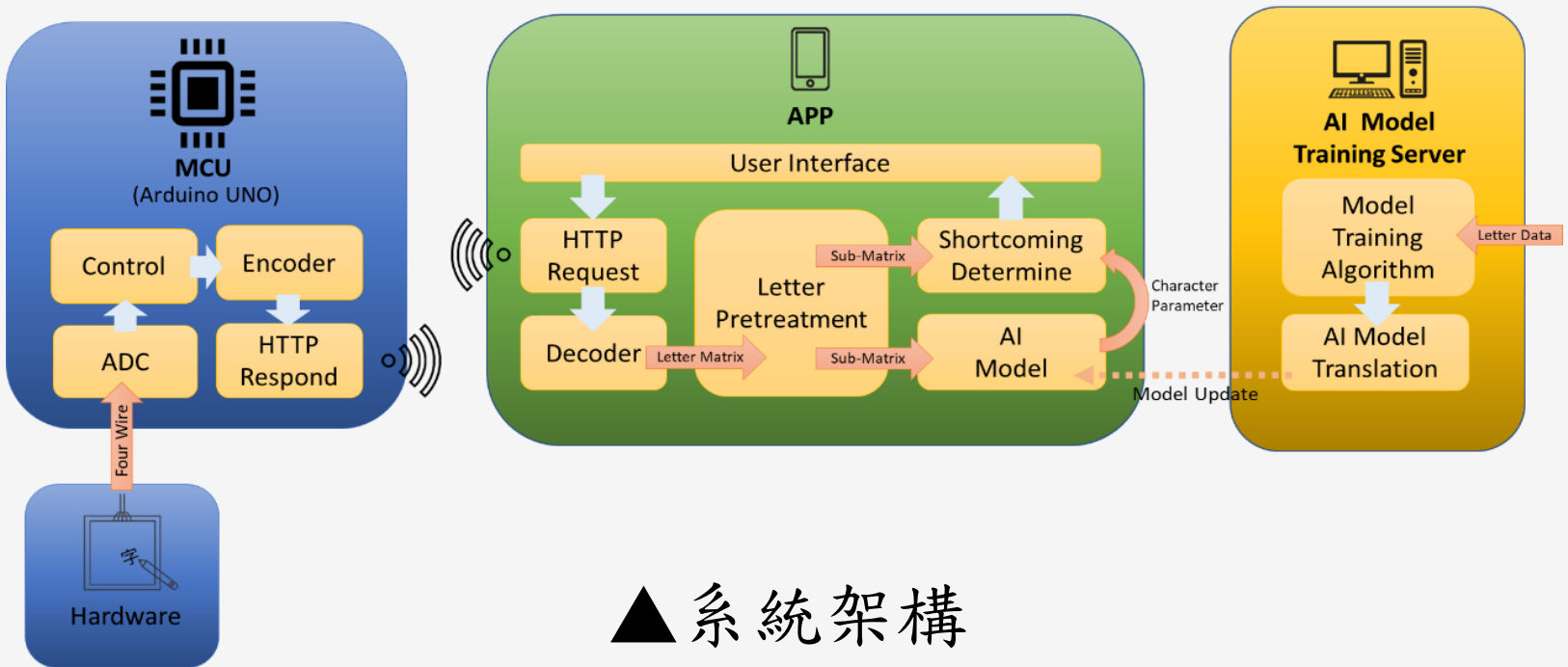


手寫字跡矯正墊板

組別：D07-1901
指導教授：陳雅淑教授

組員：劉錕笙 施丞祐 陳儀銘
學號：B10607118 B10607121 B10607133

簡介：

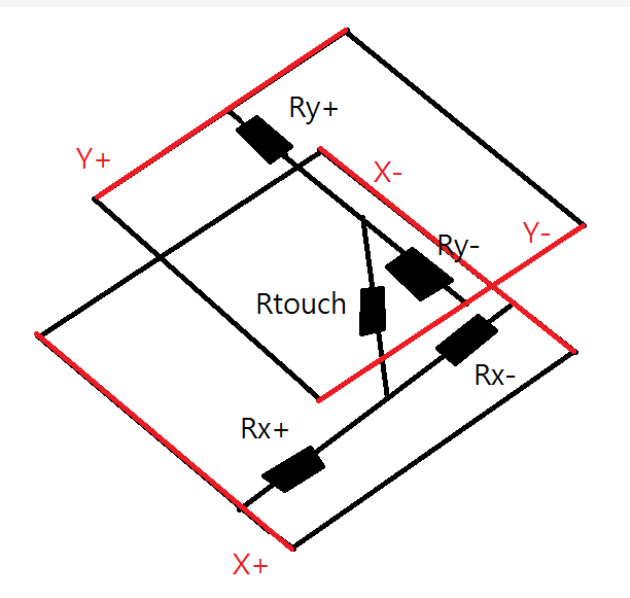


▲系統架構

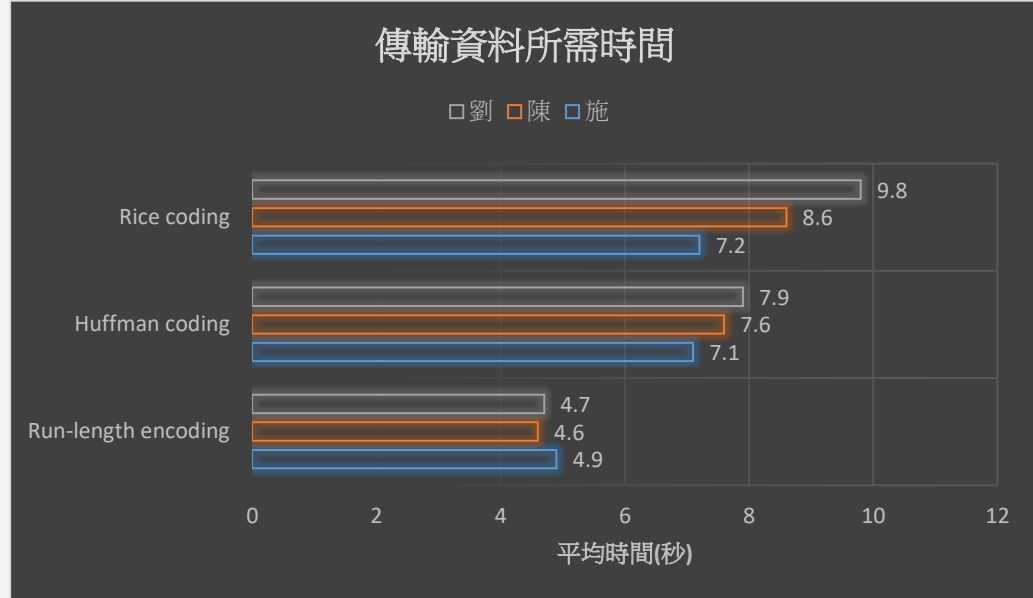
摘要：

- 提供全新練字方式
- 物件簡單，方便使用
- 壓縮傳輸，提升傳輸效率
- APP字跡評分，量化字跡美醜
- 圖形顯示回饋，改善方向明確

硬體：



▲電阻屏

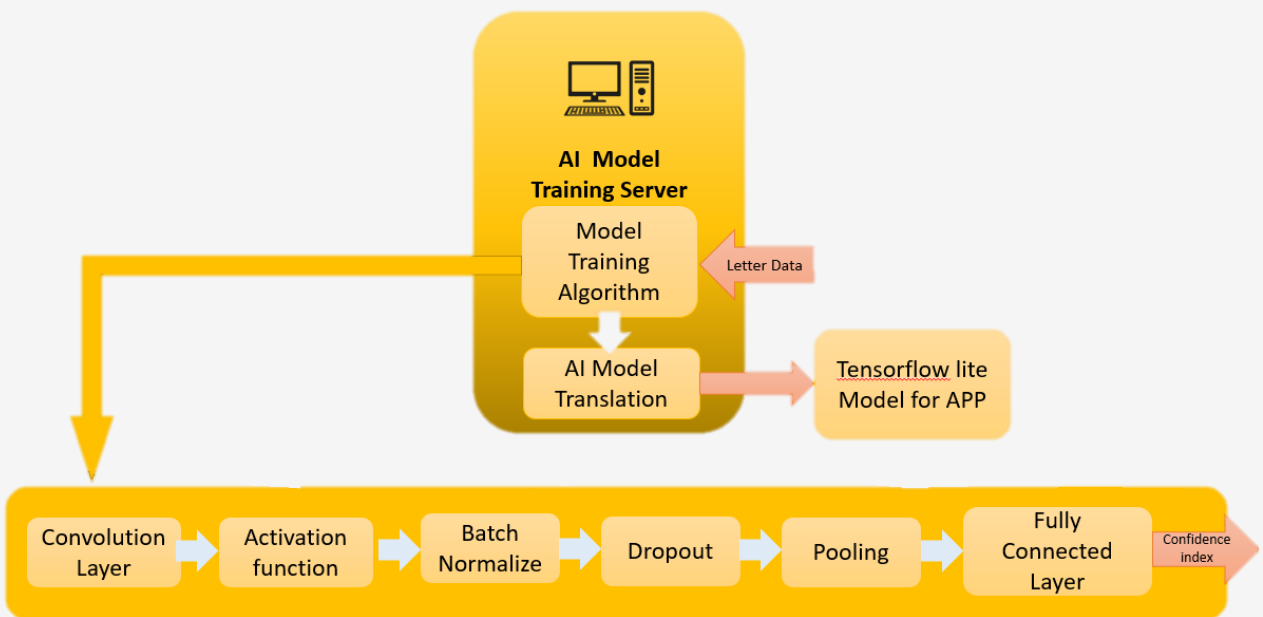


▲不同壓縮能力比較

- 使用電阻式壓力取樣作為墊板
- 使用Run-Length encoding，藉此降低傳輸壓力。
- 使用HTTP協定與APP對傳

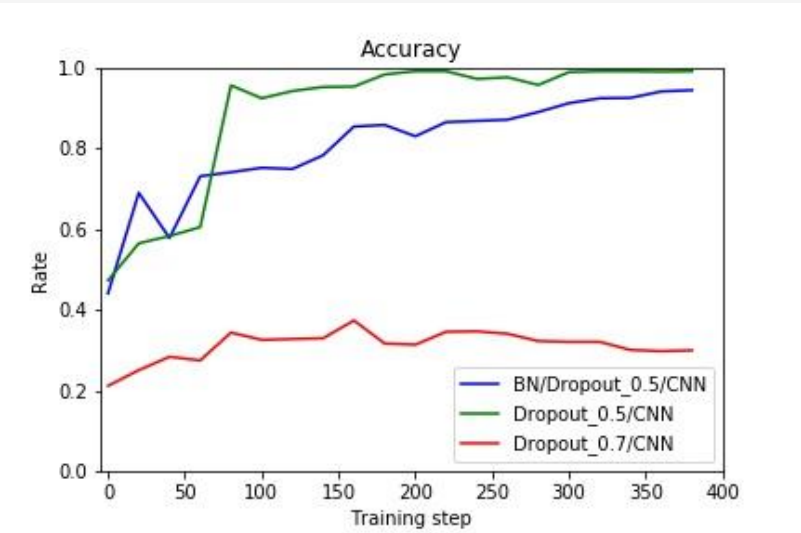
$$\frac{Data_{compressed}}{Data_{initial}} \approx \frac{\left(\frac{512^2 - n^2}{32} + n^2\right)}{512^2}, n^2 = \text{Amount of Data}$$

辨識模型：

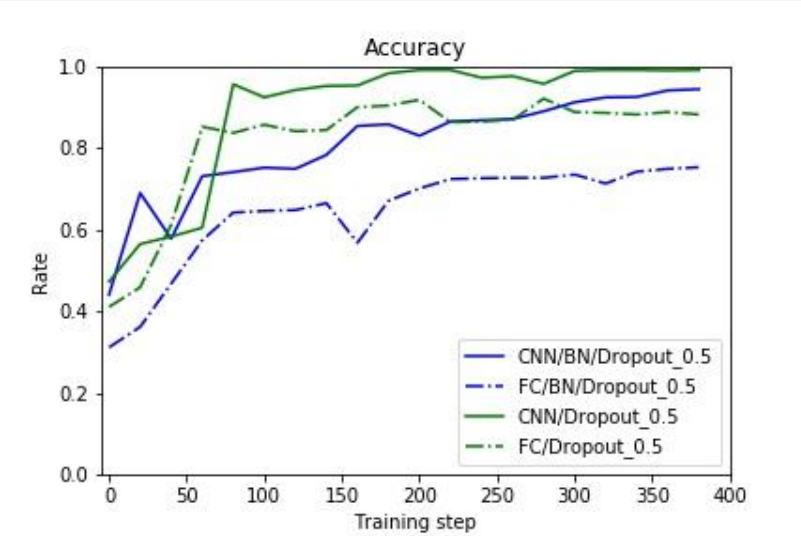


模型比較：

經由比較可得知CNN的辨識較全連接層神經網路來的好，原因是CNN使用三維的kernal對圖片進行掃描，可更有效得到資料的特徵數值，降維時也較不會損失特徵。



▲辨識率折線圖



▲捲積神經網路以及全連接層網路比較圖

```
start to train
0.1
0.1
0.1
0.1
0.1
0.1
Done
```

▲維度過高、資料量不足

```
start to train
0.6759259
0.9997487
0.9987487
0.9987487
0.9987487
0.9814815
Done
```

▲激勵函數造成過擬合

```
start to train
0.3955555
0.8240741
0.9444444
0.962963
0.962963
0.9722222
Done
```

▲改善後最終結果

說明：

第一版因為資料量不足且維度過高導致辨識率只有0.1。

說明：

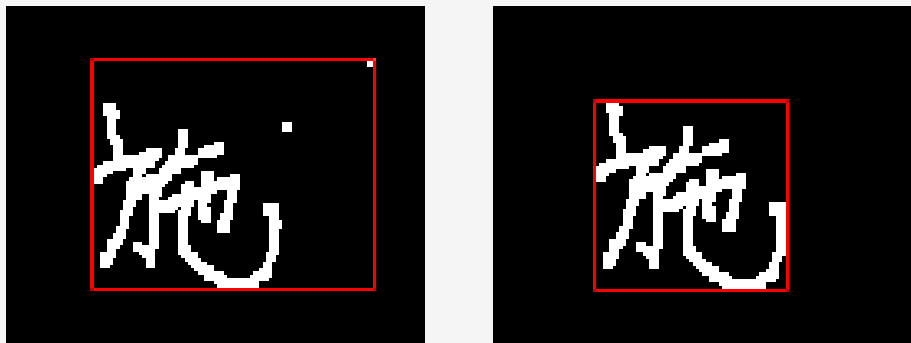
增加資料量同時降低維度至64*64，選用之激勵函數出現過擬合。

說明：

在激勵函數後加上Batch normalize使資料以常態分布改善過擬合。

字跡圈選

1. 抗躁(去雜訊)



雜訊點影響：

- 使後續演算法出錯
- 大幅增加運算量
- 使字跡失真

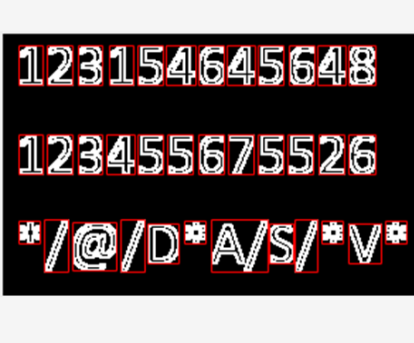
2. 斷點補償



目的：

藉由邊緣銳化(LoG)補償硬體取樣頻率不足所造成的斷點。

3. 連續字跡群組化



目的：

藉由先將字跡點分類以大幅降低後續合併運算量。並賦予字跡點字群概念。

4. 群組合併



方法：

藉由中文字方正特性設計一搜尋半徑。執行次數約為群組數的平方。

字跡回饋：



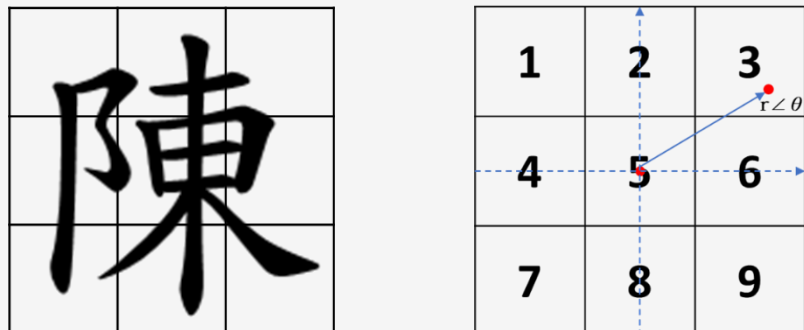
▲字跡回饋

▲改善缺陷示意圖

▲評分低

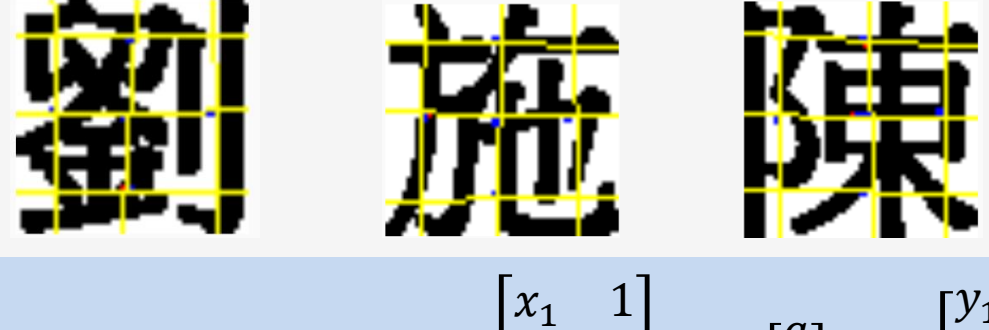
▲評分高

十點重心



採用正規化極座標，中心設置於總字跡中心，並消除長寬影響。

線性回歸



$$A^T A X = A^T B, \text{ where } A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \end{bmatrix}, X = \begin{bmatrix} a \\ b \end{bmatrix}, B = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

藉由線性回歸判斷字跡是否傾斜

結論：

評分可靠度：

藉由結構評分，未收錄字元雖使用之特徵參數不正確，但字跡評分仍具高低落差。

- 系統特性：藉由結構評分，可靠度高。
- 硬體特性：資料壓縮、傳輸效率高。
- 模型特性：CNN提升訓練效率及辨識率。