

# Deep Learning Application

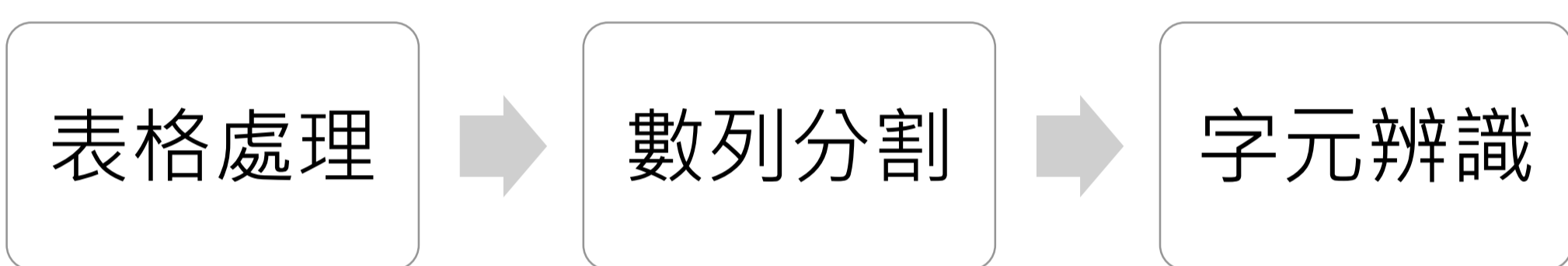
## 深度學習的應用

指導教授:劉靖家 組別: A08 組員: 游定衡、郭晉嘉、許伯謙

### 摘要

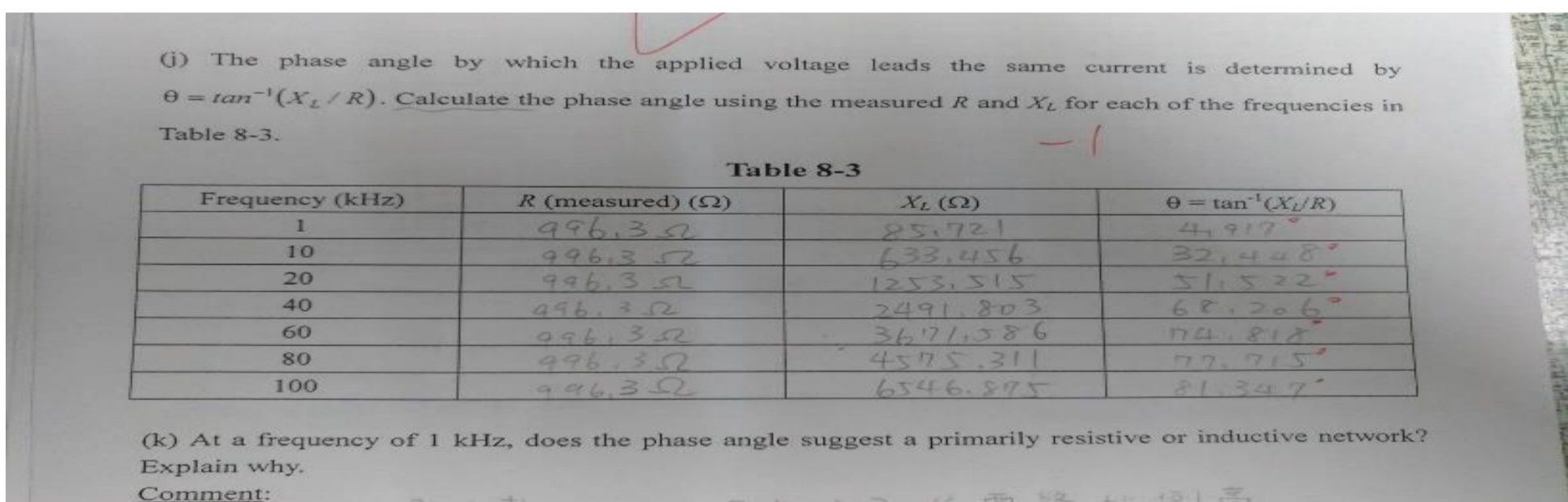
手寫的文字仍是日常生活中常見傳達與紀錄想法的媒介，要如何將手寫的文字電子化，方便搜尋、管理與分享，是一項值得探討的技術。手寫文字的辨識有許多挑戰像是文字的定位、文字的切割、文字的辨識、作者的辨識、底線偵測等。我們挑選的應用是“實驗數據表格的辨識”，希望透過影像處理和深度學習將實驗課抄寫的數據，透過電腦自動轉成Excel表格，省去人工輸入的時間，解決日常生活中會遇到的問題。

### 軟體架構



### 表格處理

(1) 讀取圖片  
設定目標資料夾的位置路徑，再讀取資料夾中所有的圖片，並以彩色圖片儲存於系統中。



(2) 表格切割  
將讀取彩色圖片轉為灰階圖片，並對圖片進行模糊處理，減少雜訊對目標資訊的影響。接著使用邊緣檢測與輪廓尋找，試著框出完整的表格邊界。此步驟會框出許多儲存格與其他判斷錯誤的邊界。因此將目標邊界設定為擁有4個頂點，且邊長為最大者，即可正確框出表格邊界。

Frequency (kHz)	R (measured) (Ω)	X <sub>L</sub> (Ω)	θ = tan <sup>-1</sup> (X <sub>L</sub> /R)
1	996.3 Ω	85.721	4.917°
10	996.3 Ω	633.456	32.448°
20	996.3 Ω	1253.515	51.522°
40	996.3 Ω	2491.803	68.206°
60	996.3 Ω	3671.586	74.818°
80	996.3 Ω	4575.311	77.715°
100	996.3 Ω	6546.875	81.347°

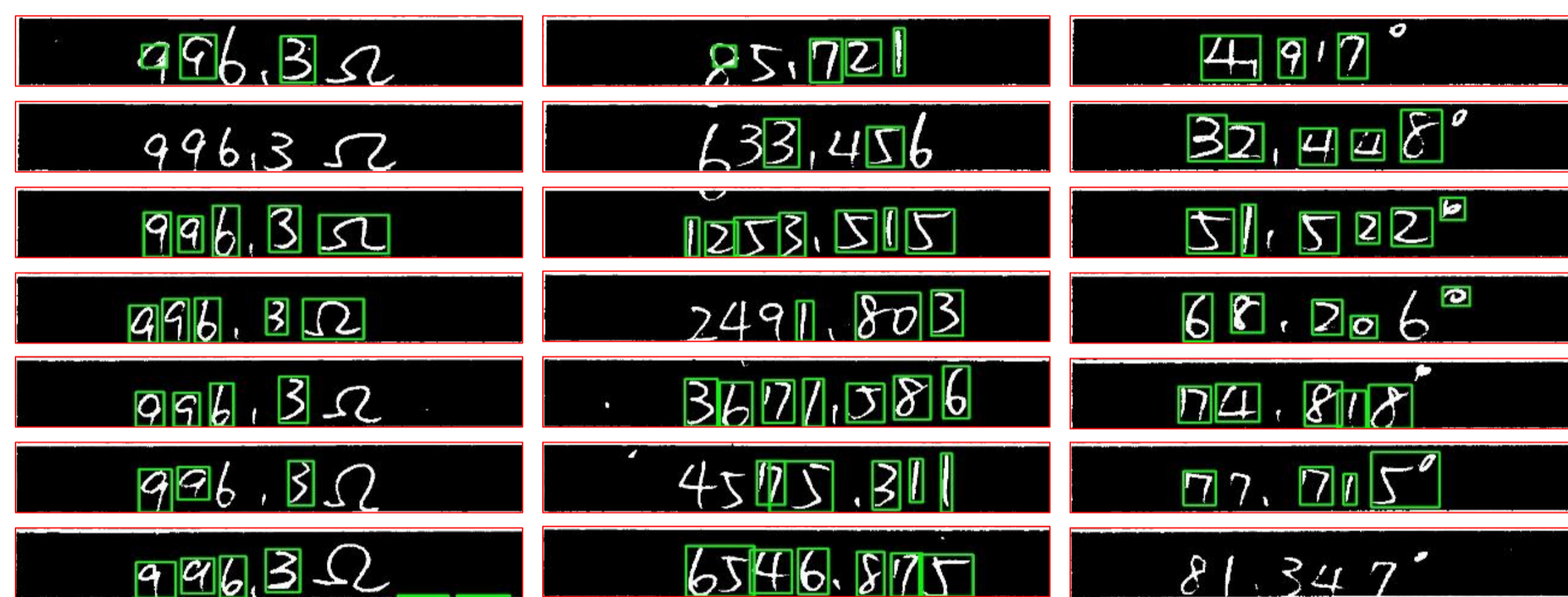
(3) 儲存格切割  
使用Hough Line Transform偵測表格中所有水平線與鉛直線，並定位其所在位置後，即可得知表格中行與列分別的間距，再藉此切割出每一個儲存格

Frequency (kHz)	R (measured) (Ω)	X <sub>L</sub> (Ω)	θ = tan <sup>-1</sup> (X <sub>L</sub> /R)
1	996.3 Ω	85.721	4.917°
10	996.3 Ω	633.456	32.448°
20	996.3 Ω	1253.515	51.522°
40	996.3 Ω	2491.803	68.206°
60	996.3 Ω	3671.586	74.818°
80	996.3 Ω	4575.311	77.715°
100	996.3 Ω	6546.875	81.347°

### 數列分割

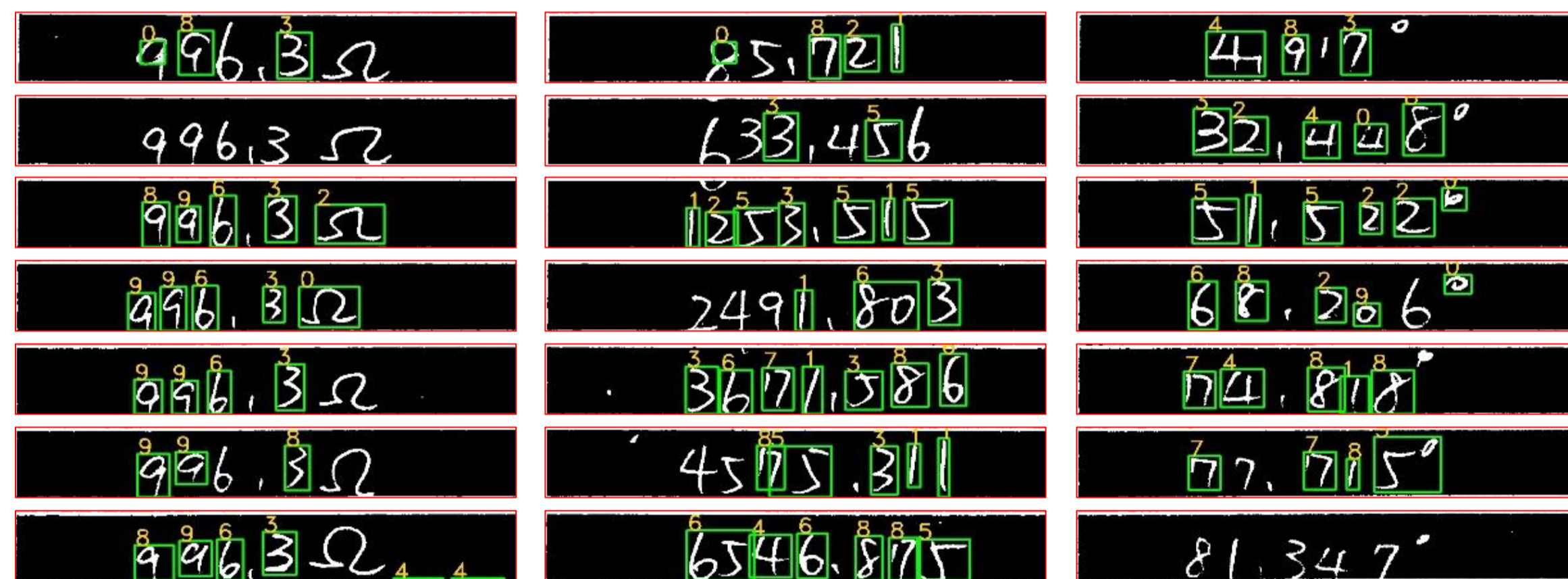
由於數列分割所關注的影像處理尺度與表格切割、儲存格切割不同，因此使用不同參數將切割下的儲存格進行模糊處理與二質化，並在二質化的過程中將顏色反轉，即深色畫素轉為白色、淺色畫素轉為黑色，以符合手寫數字辨識Dataset的規格。接著，將白色部分（手寫數字與雜訊）侵蝕，使部分雜訊消失，再將其膨脹，使手寫數字的自行更為顯著。同樣使用輪廓尋找的方式，對數字畫框，將儲存格中的一排數列分割為個別數字。並且設定數字邊框的的周長下限與上限，避免框到雜訊或判斷錯誤。

由於此CNN Model的輸入圖片必須為28×28像素的圖片，因此將切下數字的長邊設定為26像素並等比例縮放，再貼上於28×28像素的黑色背景上後，即可進入下個步驟。



### 字元辨識

事先將已使用Keras訓練完成的CNN Model匯入系統中並編譯。再將28×28像素的手寫數字圖片送進CNN Model辨識後，會將判讀的結果標示於儲存格的圖片上



### 結論

我們實作了一套系統，將相機拍攝的手寫數據表格轉成Excel檔案。包含蒐集資料、OpenCV前置處理和用Keras訓練深度學習模型。目前還有小數點、單位識別、以及字元偵測與切割的辨識率都仍須加強。OpenCV的前置處理並不具有可擴展性，需要根據不同圖片去調整參數。根據一年的研究後，未來的挑戰還有

1. 如何建立一套完整的Standard Dataset去測試不同的演算法，讓演算法不是只能適用某一種Dataset
2. 由於傳統的影像處理方式為固定的演算法，有一套計算的公式，固定的參數設定的特性，往往會受到手寫文字和相機拍攝圖片品質的不穩而導致正確率不穩。我們認為近年發展的利用深度學習做Object detection的技術是未來的方向。
3. 如何建立一套圖片的前置處理演算法，可以通用於任何相機，可以不受環境的影響保持穩定的正確率。