

科技部補助
大專學生研究計畫研究成果報告

計 畫 名 稱	: 結合多架構之七段顯示數字影像優化及辨識系統
------------	-------------------------

報 告 類 別 : 成果報告
執行計畫學生 : 葉品郁
學生計畫編號 : MOST 110-2813-C-040-070-E
研 究 期 間 : 110年07月01日至111年02月28日止，計8個月
指 導 教 授 : 秦群立

處 理 方 式 : 本計畫可公開查詢

執 行 單 位 : 中山醫學大學醫學資訊學系

中 華 民 國 111年03月25日

(一)摘要

在人們日常生活中，有許多的實驗儀器、烹飪器材或是醫療器材等設備經常會使用七段顯示器作為顯示數值的工具，例如電子秤、血壓機以及血糖機等，可以看出七段顯示器的使用頻率非常高，而以上設備的測量結果往往需要記錄下來，因此從過去到現在已有許多學者針對辨識七段顯示數字做相關研究，使用的方法從傳統的影像處理到現今蓬勃發展的機器學習都有，而最為常見的方法即為光學字元辨識(Optical Character Recognition, OCR)。然而，七段顯示數字中的不連續狀況以及欲辨識影像中的光線、角度以及數字大小等因素的不同，容易導致 OCR 辨識效果不佳。因此本計畫開發出「結合多架構之七段顯示數字影像優化及辨識系統」，利用本計畫提出之 SSC Model 將輸入影像進行反光以及角度的校正，接著 Mask R-CNN 會將七段顯示數字進行標記，利用其產生之遮罩填補七段顯示數字中不連續的狀況，其切割準確率可達 90.50%，表示 Mask R-CNN 可有效將七段顯示數字進行切割，最後使用 OCR 辨識標記後的七段顯示數字並將其顯示於本系統介面中，其中 OCR 辨識準確率為 94.40%，表示 OCR 可將 Mask R-CNN 切割後之結果轉換為數值。本計畫希望藉由大量七段顯示數字之影像訓練出能精準標記七段顯示數字之深度學習架構，以提升 OCR 辨識七段顯示數字之效能。

(二)研究動機與研究問題

由於七段顯示器之價格相較於液晶螢幕較為便宜，因此七段顯示器經常在日常生活中被用於顯示各項設備之數值，例如在進行化學實驗時用於測量物質重量以及在烹飪時用於計算食材用量的電子秤皆會用七段顯示器顯示重量公克數，而在患有慢性病的患者常用的血壓機以及血糖機中亦會使用七段顯示器顯示測量之結果，且以上數值經常需要進行紀錄，以作為後續分析或醫療診斷之使用，因此有許多民眾會藉由手抄的方式將結果進行紙本記錄，然而此方式容易產生謄寫錯誤的狀況，也會造成民眾記錄數值的不便。雖然目前已有光學字元辨識可作為辨識工具，且市面上已有許多軟體提供該項功能，但由於七段顯示數字間不連續，如圖 1 所示，與一般常見的印刷體或書寫體皆有較大的差異，使得 OCR 技術應用於七段顯示數字之辨識效果普遍不佳。此外，OCR 起初是應用於文本掃描，以將紙本文件電子化，因此對於文件拍攝時的光線、角度以及距離非常講究。

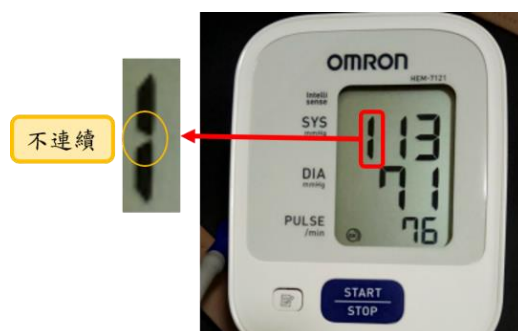


圖 1、七段顯示器數值間不連續

近年來深度學習蓬勃發展，其能在接收到大量的資料後自行歸納成有用的資訊，並針對輸入的影像進行分析或分類。在過去，已有諸多學者提出不同影像處理之方法應用於辨識數字，如今更是發展出各式深度學習演算法在手寫數字辨識以及車牌數字辨識方面，並且實際應用於日常生活中。然而，由於七段顯示器數字間不連續的情形，提高了辨識七段顯示數字的門檻。因此，本計畫希望結合 SSC Model、Mask R-CNN 提升 OCR 辨識七段顯示數字的成效。透過 SSC Model (Specular-Skew Correction Model) 中的 Deep k-Means 期望其能將輸入影像分為有反光且有傾斜、有反光且無傾斜、無反光且有傾斜以及無反光且無傾斜共四群，並將其中有反光的兩群影像利用 Specular Correction Model 去除影像中反光的白點，接著將有傾斜的兩群影像輸入至 Skew Correction Model 以校正影像的歪斜角度，再透過 Mask R-CNN 標記影像中的七段顯示數字，以此將不連續的區域進行填補，最後利用 OCR 辨識標記後的七段顯示數字。

(三)文獻回顧與探討

在人們日常生活中，處處皆有七段顯示器的存在，例如電子錶、電梯樓層以及血糖機等，而在各項用來測量某物質含量的儀器上，更是經常使用七段顯示器顯示測量結果，使用頻率非常高且往往需將測量結果記錄下來，因此不論是過去傳統的影像處理亦或是現今蓬勃發展的深度學習都有許多學者針對其進行相關研究，而本計畫將參考多篇文獻作為基礎，以發展出能提升七段顯示數字辨識效果的方法，提高人們日常生活的便利性。

1. 深度學習應用於去除影像中反光狀況之相關研究

由於影像中反光的白點容易造成影像中資訊的缺失，因此 Madessa 學者等人利用 Partial Convolution Layer[1]可進行影像修復的特性，將其應用於具有鏡面反射或反光的影像中，以此將影像中的亮光區域進行填補，且其訓練影像包含多樣化的影像，如各式物品之影像、醫學影像以及湖面影像等，故其提出之架構可應用於各式影像中。Lin 學者等人提出利用 CNN[2]抽取影像中具有高亮度的區域，並將其根據周圍的像素值進行填補，並利用 GAN 生成對抗網路中的鑑別器分辨原始影像以及處理後之影像，以確保處理後之影像前後一致，但由於此方法會根據影像中的暗處進行填補，因此容易使整張影像偏暗。

2. 深度學習應用於影像傾斜校正之相關研究

由於影像傾斜容易影響辨識數字或字元的效果，因此 Yoo 學者等人針對車牌辨識提出一種新的 DNN 架構[3]，用以預測出傾斜之車牌經過校正後的四個端點座標位置，再利用 Affine Transform 進行傾斜校正，且證明此方法能有效提升車牌辨識之辨識率。Akhter 學者等人針對文本歪斜的狀況進行校正，其首先利用 ConvNet 進行影像之歪斜偵測，以確認影像歪斜之角度，再利用其提出的深度學習架構[4]進行校正，然而，此方法僅針對歪斜狀況進行校正，若從側面拍攝目標物，則影像產生之傾斜狀況將無法進行校正。

3. 傳統影像處理方法應用於七段顯示數字辨識之相關研究

由於在背景複雜的七段顯示器影像中難以偵測與辨識數字或字元，因此 Popayorm 學者等人針對紅色的 LED 七段顯示器進行研究，提出預定義 HSV 顏色切片技術 (predefined HSV color slicing technique)[5]，用以在影像中準確地找出數字及字元的邊緣，並將其切割出來。此論文所提出的方法可以將影像的前景與背景分離，以提升辨

識 LED 七段顯示器數字與字元的表現。Kanagarathinam 學者等人為了辨識出數位電錶上的數字，因此他們所採用的「YUVAEB 資料集」中包含傾斜角度與亮度不同的七段顯示器影像，之後使用 MSER 方法進行文本檢測，並將其進行二值化以及膨脹等影像處理，最後透過 OCR 進行數字辨識[6]。然而，此一連串的辨識以及處理過程非常耗時。同年，Tsiktsiris 學者等人則提出了加速光學字元辨識七段顯示器數字的方法[7]，以實現自動記錄血壓計與血糖機的測量結果，輔助經常使用醫療設備的年長者，但由於 OCR 為文本掃描的工具，因此較難以針對過於歪斜的影像進行辨識。

4. 深度學習應用於各種數字偵測及辨識之相關研究

由於在鋼鐵製造業中，煉鋼所需的材料—鋼坯具有不同的成分、厚度及長度等差異，因此鋼坯上會刻有識別號(BIN)，用以得知該鋼坯的相關資訊，而 Koo 學者為了成功辨識鋼坯上的 BIN，提出了一個深度學習網路框架[8]，其可以同時偵測與辨識 BIN，且當影像進行旋轉時，仍能成功識別 BIN，然而，BIN 的字型以及拍攝背景皆與一般所熟知印刷體、七段顯示數字以及明亮背景有所不同，因此較難以應用於日常生活中。Kalyan 學者等人提出了一個 Android 應用程式，能在智慧型手機等移動設備上，利用 Inception V3 深度學習模型進行七段顯示器的數字辨識[9]，證明深度學習亦可應用於七段顯示數字的辨識中。

5. Mask R-CNN 應用於物件切割之相關研究

為了辨識出指針式儀表中所顯示的數值，Lin 學者等人提出利用 Mask R-CNN[10]將指針式儀表中的指針標記出來，並利用其提出的 PRM 演算法計算指針所指向的數值，以此自動讀取指針數值，證明了 Mask R-CNN 可以精準切割出物件的特性。而 Seimi 學者等人為了進行車牌辨識，則利用 Mask R-CNN[11]將汽車影像中的車牌切割出來，接著將切割後的車牌影像再次輸入至 Mask R-CNN 中，用以對車牌中的數字、字母及字元進行標記以及分類，達到辨識車牌的目的，證明了 Mask R-CNN 除了物件標記及切割外，亦可用於數字辨識。

綜合上述文獻，本計畫發現目前並未有學者利用深度學習優化七段顯示數字影像，以提升辨識效果，但有許多方法在影像的反光問題以及傾斜問題上都可達到良好的結果。因此本計畫希望結合上述文獻所提出的想法或方法，並搭配自行建置的圖形化介面，開發出「結合多架構之七段顯示數字影像優化及辨識系統」，期望透過此系統能提升人們在讀取七段顯示數字時的便利性。

(四)研究方法及步驟

本計畫結合 SSC Model、Mask R-CNN 以及 OCR 建置出「結合多架構之七段顯示數字影像優化及辨識系統」，本系統流程圖如圖 2 所示，其主要分為四個部分，分別為輸入七段顯示器影像、SSC Model 影像校正、Mask R-CNN 切割及標記七段顯示數字模型以及光學字元辨識七段顯示數值，以下將針對各部分詳細說明。



圖 2、系統流程圖

1. 輸入七段顯示器影像

本計畫利用智慧型手機拍攝各式七段顯示器影像，其中包含血壓計影像以及電子秤影像，共有 1103 張影像做為訓練 SSC Model 之訓練資料集，此資料集的每張影像中皆以七段顯示數值顯示儀器測量之結果。為了增加影像的多樣性，本計畫也將收集到的影像進行角度及亮度的調整，在角度調整的部分，分別將影像旋轉 0 度、90 度、180 度以及 270 度，而在亮度調整的部分，則是分別將影像調亮 30% 以及調暗 30%，經過資料擴增後，總共有 13236 張，如圖 3 所示，而其中 90% 的影像作為訓練影像，10% 作為測試影像。



圖 3、模型訓練資料集影像，
(a) 為未經處理之血壓計影像，(b) 為未經處理之電子秤影像

2. SSC Model 影像校正

本計畫使用 SSC Model 校正影像中的反光以及傾斜狀況，其中使用 Deep k-Means 將輸入的七段顯示數字影像分為有反光且有傾斜、有反光但無傾斜、無反光但有傾斜以及無反光且無傾斜共四群。SSC Model 中包含三個部分，分別為 Deep k-Means、Specular Correction Model 以及 Skew Correction Model。SSC Model 流程圖如圖 4 所示。

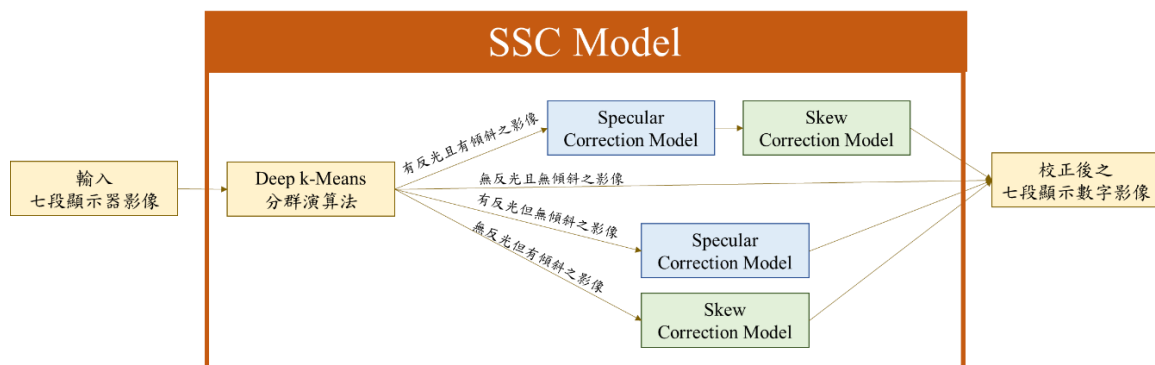


圖 4、SSC Model 流程圖

2.1 Deep k-Means 分群演算法

為了依照不同影像中的情況進行不同的處理，以縮短系統處理時間，因此本計畫將含有七段顯示數字之影像輸入至 Deep k-Means 分群演算法中，以此針對影像的傾斜以及反光狀況進行分群，共可分為四群，分別為有反光且有傾斜的影像、有反光但無傾斜的影像、無反光但有傾斜的影像以及無反光且無傾斜的影像。其結合了自動編碼器 (Auto-Encoder) 以及 k-Means 分群法，利用深度學習具有學習的特性提升傳統 k-Means 分群法的分群效果。

2.2 Specular Correction Model 校正影像中的反光狀況

由於七段顯示器之螢幕表面通常較為光滑，因此可能受到周圍光線干擾而有反光的情況發生，而為了去除影像中的反光白點，本計畫會將分群後的影像中有反光且有傾斜的影像以及有反光且無傾斜的影像皆輸入至 Specular Correction Model 中，以此去除影像中反光的白點。Specular Correction Model 中採用經常用於影像修復的 Partial Convolutional Layer (Pconv)來抽取原始影像中的特徵。其會先將輸入之七段顯示數字影像進行二值化，以此找出影像中的反光區域，並將此二值化影像作為遮罩，黑色標示非反光區域，白色表示反光區域，再將此遮罩以及原始七段顯示數字影像一同輸入至 Pconv 以此對照遮罩區域來抽取原始影像中的特徵，填補影像的反光區域。Specular Correction Model 架構圖如圖 5 所示。

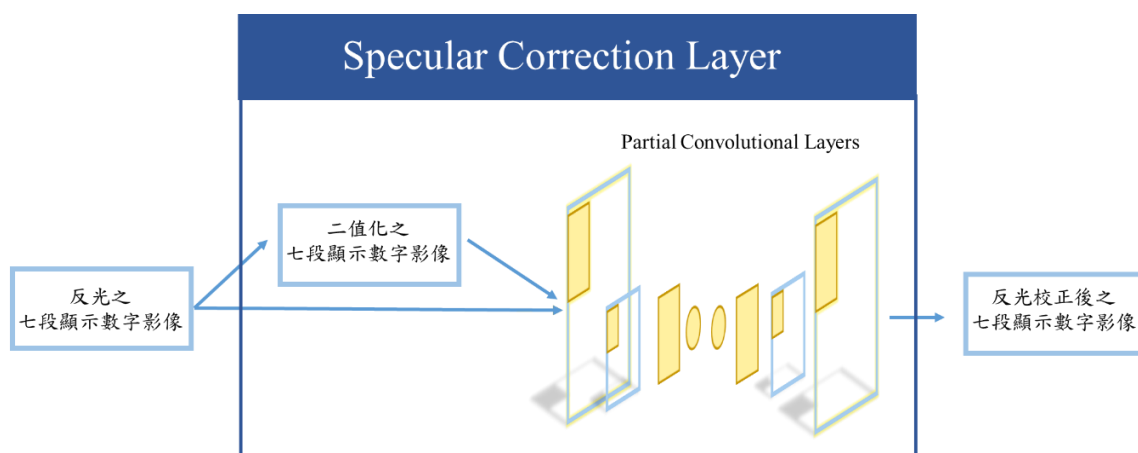


圖 5、Specular Correction Model 架構圖

2.3 Skew Correction Model 校正影像的傾斜狀況

為了避免因影像傾斜而導致後續之數字切割及辨識效果不佳，本計畫會將分群後的影像中無反光且有傾斜的影像以及經 Specular Correction Model 校正後之有反光且有傾斜的影像輸入至 Skew Correction Model 中，其架構含有三層卷積層以抽取七段顯示數字影像中的特徵以及三層池化層來縮減卷積層抽出之特徵圖大小，接著會利用三層全連接層輸出七段顯示區域校正後之四個座標點，最後會根據先前預測之新座標點進行 Affine Transform，以此校正影像的傾斜狀況。Skew Correction Model 架構圖如圖 6 所示。

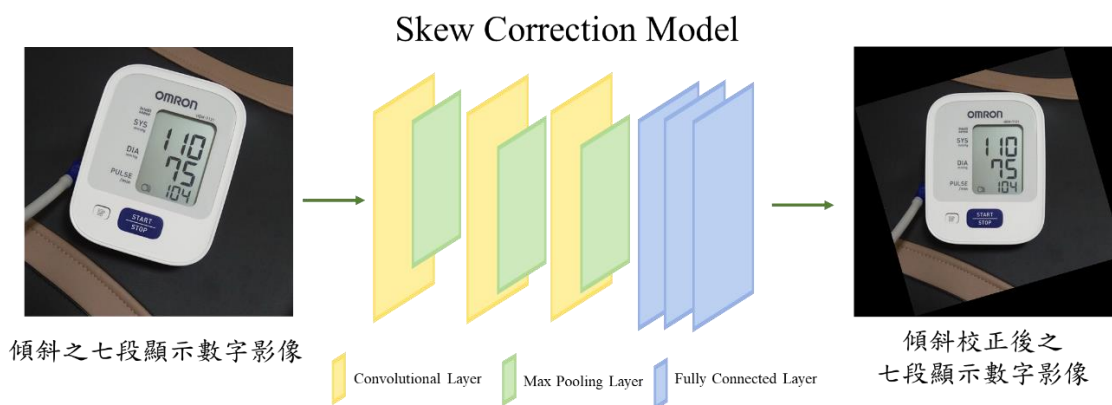


圖 6、Skew Correction Model 架構圖

3. Mask R-CNN 切割及標記七段顯示數值模型

由於七段顯示數字中含有不連續情形，會導致常見的 OCR 技術無法辨識，因此本計畫希望透過具有物件切割功能的深度學習模型，利用其所生成的遮罩，將七段顯示數字中的不連續情形進行填補。而 Mask R-CNN[10-11]會針對影像的每個像素進行標記，相較於其他切割演算法更能達到精準標記物件的效果，因此本計畫會將經過 SSC Model 處理後之七段顯示數字影像輸入至 Mask R-CNN 模型中進行標記和分類，Mask R-CNN 架構圖如圖 7 所示，首先會先經過由 ResNet101 及 FPN(Feature Pyramid Networks)所構成的 Backbone 中抽取七段顯示數字影像中的特徵並輸出特徵圖，再將此特徵圖及七段顯示數字影像輸入至 RPN(Region Proposal Network)中，偵測七段顯示數字的可能位置。接著，Mask R-CNN 會將偵測到可能含有七段顯示數字的位置及其對應的特徵圖輸入至 RoI(Region of Interest) Align 中進行調整，並將調整後的結果輸入至全連接層進行分類及遮罩的定位，最終，Mask R-CNN 會將影像中的數字依照其外型分別進行標記以及給予其相應之類別和遮罩顏色。

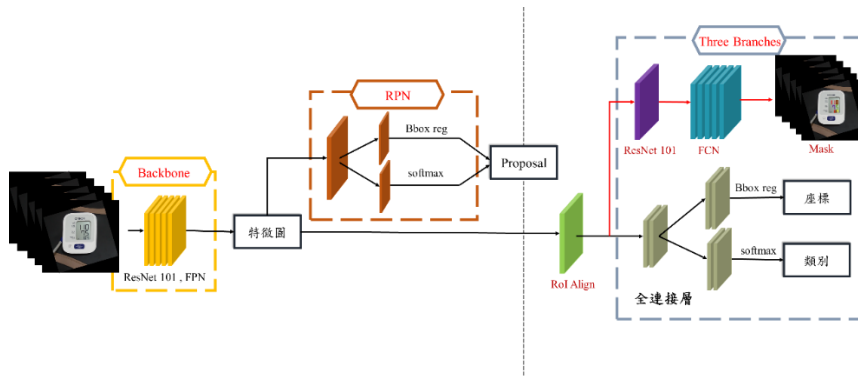


圖 7、Mask R-CNN 架構圖



圖 8、七段顯示數字之標記影像

為了評估 Mask R-CNN 模型的訓練成效，本計畫使用了 RPN loss、RoI loss、class loss、bounding box regression loss 以及 mask loss，共五個損失函數。其中，RPN loss 是用於計算 proposal 的 classification loss 以及其 bounding box loss 的總和，用以進行位置的調整，RPN loss 如公式(1)所示。而 RoI loss 則是用來計算最終輸出的類別的 classification loss、bounding box regression loss 以及遮罩 mask loss 的總和，以求得模型的最佳切割效能，RoI loss 如公式(2)所示。

$$L_{rpn} = L_{class} + L_{bbox} \quad (1)$$

$$L_{RoI} = L_{class} + L_{bbox} + L_{mask} \quad (2)$$

其中 L_{class} 表示 class loss， L_{bbox} 為 bounding box loss， L_{mask} 為輸出遮罩的 loss。

4. 光學字元辨識

為了辨識影像中的七段顯示數值，本論文採用 Tesseract 光學字元辨識技術作為辨識方法[12]。由於 Tesseract-OCR 是由大量且多元的文字影像訓練而成的辨識引擎，其訓練影像也包含七段顯示數字、中文字及標點符號等，因此 Tesseract-OCR 在七段顯示數字的辨識方面相較於傳統 OCR 優異。本論文會先將 Mask R-CNN 切割並標記後的影像輸入至 OCR 中進行辨識。而 OCR 會透過垂直投影以及水平投影的方式，取得影像中各物件的位置，以此將影像中的文字、數字及標點符號切割出來，再對切割出的每一個物件進行侵蝕，用以在保留重要訊息的同時，亦可將雜訊或其餘的資料去除，避免在辨識時造成誤判，接著，會利用統計的方式抽取出影像中物件的特徵，並將其與資料庫中的字元進行比對，最後輸出具有最高相似度的字，以此完成七段顯示數值的辨識。

(五) 實驗結果與探討

在實驗結果與探討的部分，首先將介紹系統開發環境與使用的工具，接著針對實驗結果計算混淆矩陣及相關的評估指標，並根據實驗結果進行系統的分析與討論，最後介紹系統使用者介面。

1. 開發環境與使用工具

本計畫在訓練及測試階段皆使用相同開發環境與設備，軟硬體設備配置如表 1 所示。本計畫於 Windows 10 作業系統中使用 Python 程式語言搭配 Keras 框架進行開發。硬體設備使用 Intel Core i9-9900X、NVIDIA Geforce RTX 2080 Ti 和 64GB RAM 的電腦。

表 1、軟硬體設備

軟體	作業系統	Windows 10
	開發程式語言	Python
	深度學習框架	Keras
	GUI 套件	TKinter
硬體	CPU 處理器	Intel Core i9-9900X
	GPU 顯卡	NVIDIA Geforce RTX 2080 Ti
	RAM 容量	64G

2. 結果討論

為了評估 Mask R-CNN 深度學習方法應用於七段顯示數字的成效，本研究透過計算混淆矩陣的相關指標來做為本實驗結果的評估指標，並將 Mask R-CNN 的結果分為切割以及分類兩部分進行探討。根據 Mask R-CNN 切割七段顯示數字之混淆矩陣所計算出的評估指標，測試準確率為 90.50%，得以證明 Mask R-CNN 能夠準確的切割出數字區域，且其分類正確；在精確度的部分則有 90.94%，表示 Mask R-CNN 標示七段顯示數字的位置及類別大多正確；測試敏感度為 91.86%，表示 Mask R-CNN 對於七段顯示數字的區域有高靈敏度因此可正確將其進行切割；最後，特異度在測試階段為 88.96%，表示實際為非七段顯示數字的區域中，正確被切割的比例高達 90% 以上。由此可知，Mask R-CNN 切割七段顯示數字具有相當不錯的成效，證明本研究提出的方法能夠準確標記出七段顯示數字位置，以填補七段顯示數字中不連續的狀況，如圖 9

所示。然而從 Mask R-CNN 切割結果圖中可以發現，Mask R-CNN 大多可以將七段顯示數字的位置準確切割出來，但容易有類別辨識錯誤或類別重疊的問題，例如圖 9 下排第一行中的數字 6，雖然能被完整切割出來，但其類別被辨識為數字 8；而圖 9 下排第二行中的數字 7，同樣能被完整切割出數字 7 的外型，且類別標示正確，但同時也被辨識為數字 1。



圖 1、Mask R-CNN 切割七段顯示數字之結果

因此我們將進一步利用混淆矩陣探討 Mask R-CNN 的分類成效，從混淆矩陣所計算出的評估指標，測試準確率為 68.71%，平均測試精準度為 85.12%，平均測試敏感度為 53.68%，平均測試特異度為 68.39%，由上述評估指標可以得知若單獨使用 Mask R-CNN 作為辨識七段顯示數字的方法其效果不佳，其中我們也發現樣本數愈多，其敏感度愈高，因此愈容易將類似外型的數字辨識錯誤，例如圖 9 下排第二行中的數字 6，因為在七段顯示中與數字 8 相似，因此容易誤判，進而導致精準度愈低。而為了增強七段顯示數字辨識的成效，本計畫在 Mask R-CNN 切割七段顯示數字後，將產生之遮罩區域進行二值化，以此取得各類別遮罩的聯集，再利用光學字元辨識技術對此二值化影像進行數字辨識，辨識成效如表 2 所示。由下表可發現，Mask R-CNN 搭配光學字元辨識技術之數字辨識在各項評估指標中都可提升約 10%。

表 2、七段顯示數字辨識成效

	準確率	精準度	特異度	敏感度
僅使用 Mask R-CNN 之辨識結果	68.71%	85.12%	53.68%	68.39%
使用 Mask R-CNN 及光學字元辨識之辨識結果	94.40%	95.09%	90.19%	97.22%

3. 系統實作與介面展示

本計畫之系統的介面是使用 Tkinter 函式庫進行開發，並以簡潔易懂的介面為開發目標，期望有效加速醫師診斷時間。系統主介面如圖 10(a)所示，使用者點選「開啟檔案」將拍攝到的七段顯示數值影像輸入至本系統中，接著點選系統左側選單中的「分析結果」，系統會將影像進行校正，並且輸入至 Mask R-CNN 模型中將影像中的數字切割出來，如圖 10(b)所示，並記錄辨識後的數值，以避免使用者因手抄數值而產生錯誤，提升使用者日常生活的便利性。



(a)



(b)

圖 10、系統介面圖，(a)為輸入欲辨識之影像後之介面，
(b)為辨識七段顯示數字結果之介面

(六) 結論

隨著慢性病患者的人口逐年增加，使用醫療設備測量生命表徵的人數也越來越多，而七段顯示數字為目前醫療器材常用於顯示測量結果的顯示方式，但由於七段顯示數字中的不連續狀況，使得病患無法利用目前市面上已有的辨識工具進行辨識及記錄，因此本計畫提出利用 SSC Model 自動校正影像中的反光以及傾斜角度，並根據 Mask R-CNN 深度學習方法自動標記影像中七段顯示數字能達到完整切出數字的部分，再使用 OCR 辨識標記後的影像，如此能增加數字的辨識成功率。在未來，我們期望能改善及縮減影像前處理的步驟，藉以降低整體辨識所花費的時間。

(七) 參考文獻

- [1] A. H. Madessa, J. Dong, Y. Gan, and F. Gao, "A deep learning approach for specular highlight removal from transmissive materials," *Expert Systems*, 2020.
- [2] J. Lin, M. E. A. Seddik, M. Tamaazousti, Y. Tamaazousti, and A. Bartoli, "Deep Multi-class Adversarial Specularity Removal," *Scandinavian Conference on Image Analysis*, pp. 3-15, 2019.
- [3] H. Yoo, and K. Jun, "Deep corner prediction to rectify tilted license plate images," *Multimedia Systems*, 2020.
- [4] S. S. M. N. Akhter, and P. P. Rege, "IMPROVING SKEW DETECTION AND CORRECTION IN DIFFERENT DOCUMENT IMAGES USING A DEEP LEARNING APPROACH," *International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1-6, 2020.
- [5] S. Popayorm, T. Titijaronroj, T. Phoka and W. Massagram, "Seven Segment Display Detection and Recognition using Predefined HSV Color Slicing Technique," *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2019.
- [6] K. Kanagarathinam, K. Sekar, "Text detection and recognition in raw image dataset of seven segment digital energy meter display," *Energy Reports*, Vol. 5, pp. 842-852, 2019.
- [7] D. Tsiktsiris, K. Kechagias, M. Dasygenis, P. Angelidis, "Accelerated Seven Segment Optical Character Recognition Algorithm," *IEEE 2019 Panhellenic Conference on Electronics and Telecommunications (PACET)*, 2019.
- [8] G. Koo, J. Yun, H. Choi, S. Kim, "Unified deep neural networks for end-to-end

- recognition of multi-oriented billet identification number,” *Expert Systems with Applications*, Vol. 168, 2021.
- [9] V. Kalyan, M. Katti, S. Pai, “Application of Convolutional Neural Networks in Mobile Devices for Inferring Readings from Medical Apparatus,” *International Journal of Research and Scientific Innovation (IJRSI)*, Vol. 4, pp. 49-52, 2017.
- [10] Z. Lin, P. He, C. Zhang, Z. Zhang, “A robust approach to reading recognition of pointer meters based on improved mask-RCNN,” *Neurocomputing*, Vol. 388, pp. 90-101, 2020.
- [11] Z. Seimi, M. B. Halima, U. Pal, M. A. Alimi, “DELP-DAR system for license plate detection and recognition,” *Pattern Recognition Letters*, Vol. 129, pp. 213-223, 2020.
- [12] E. Zacharias, M. Teuchler, B. Bernier, “Image Processing Based Scene-Text Detection and Recognition with Tesseract,” *Computer Vision and Pattern Recognition*, Vol. 2004, 2020.