

科技部補助
大專學生研究計畫研究成果報告

計 畫 名 稱	： 人工智慧混合型深度學習網路應用於心臟醫學影像之組織分割與建模
------------	----------------------------------

執行計畫學生：張書榕

學生計畫編號：MOST 109-2813-C-008-045-E

研究期間：109年07月01日至110年02月28日止，計8個月

指導教授：陳健章

處理方式：本計畫可公開查詢

執行單位：國立中央大學生醫科學與工程學系

中華民國 110年02月19日

人工智慧混合型深度學習網路

應用於心臟醫學影像之組織分割與建模

I. 摘要

心臟核磁共振是檢查心臟疾病的重要指標之一，醫生除了可以藉由影像判斷組織有無病變，亦可從連續跳動的行為模式獲得重要的生理檢驗指標，如各心房心室容積、射血分數(Ejection fraction)等等。但若想從時序影像上獲得上述參數則需要專家花費大量時間標記外框與計算。因此本計畫透過混合型的深度學習模型幫助醫生有效率的計算出各時序下的心臟容積，並協助求出各項重要的醫學指標參數，同時期望以此技術做為基礎朝不同造影方法的醫學圖像邁進。

II. 研究目的

根據衛福部報告指出心臟疾病長期佔據了國人十大死因中的前三名[1]，也因此市面上浮現許多偵測心臟變化的儀器，如智慧型手環可供使用者參考心臟健康程度，但更深入的資訊我們卻無從得知。因此醫生會建議患者進行各項深入的檢驗如超音波、CT、MRI 等等。超音波成像技術雖可看出心臟的粗淺輪廓但卻難以判斷微血管的受損程度[2]；CT 雖能獲得微血管和鈣化程度但仍無法精確判斷心肌纖維化的程度[2]；幸好 MRI 成像技術除可精確獲得心肌病變情形亦能便於醫生釐清病人心臟血液流動狀況。由以上所述可知，深入的 MRI 檢驗是項很好的非侵入檢驗方式，因此國際醫學組織認為心臟 MRI 是判斷心臟疾病的黃金指標。

現今深度學習技術正趨成熟，國內外亦有很多研究團隊致力於心臟影像的辨識與切割，同時更有許多公開競賽如 MICCAI、Kaggle 的挑戰等吸引研究者目光。然而心臟 MRI 在各項挑戰的正確率卻仍不甚理想，其原因在於心肌與周圍臟器的對比度較低、充血與缺血時心臟對比度不一、掃描時分辨率低以及運動偽影等難以獲得解決的工程問題[3]，因此針對心臟 MRI 精準定位以及精準邊界切割仍是該領域中最值得深思與探討的問題。該問題若能獲得充分解決，則在臨床研究以及臨床判斷上將能替醫生或技師提供更為豐富且有意義的資訊。

本計畫中，我們將醫學影像切割常用的語意分割技術 U-Net[4]、可在通道中學習之注意力模型 Selective Kernel Networks (SK-Net)[5]，和 SK-Net 同為注意力模型架構的 Convolution Block Attention Module (CBAM)[6]等三種技術，系統性的融合成本計畫中所採取的心臟影像辨識與分割策略。希望擷取各技術之優點，而能在動態的心臟 MRI 下更精確算出左心室容積量，以作為臨床上判斷心臟疾病嚴重與否的指標，並希望能在較低階之硬體上實現此方法，以便未來能移植到醫院的系統。

III. 文獻回顧

現今語意分割技術主要以 Fully Convolutional Network(FCN)[7]和 U-Net 為主流技術。FCN 在編碼階段用一系列的捲積(convolutions)與池化層(pooling layers)擷取出圖像特徵，解碼時再以上採樣(upsampling)或反捲積層(deconvolutions)來還原圖像之大小；

而 U-Net 則由 FCN 變化而來，其架構在編碼時會保留前期特徵，待解碼時會與上採樣之特徵連接，因此可以加重深層並保留淺層訊息，這項特點也使得 U-Net 適合用於小資料集的訓練。

在歷年競賽與論文中有許多研究團隊依此兩種技術進行改良，其方法也十分多樣化。例如文獻[8]將其加上空間與時間上之訊息，因而提出以多面向之 MRI 影像來學習心臟形狀先驗資訊，其結果也證實多面向的計算可優於一般的 2D U-Net。但亦並非每個資料集都有其他面向的資訊供人參考，因此亦有研究團隊提出使用多尺度的特徵向量來增強 FCN 或 U-Net 之效果。例如文獻[9]所提出的 Dense-Unet 可將不同大小之特徵向量融合使得在高變異度之心臟 MRI 中得到較優異的成績，但其仍舊無法克服心尖與底部切割不良之後果。

除了架構上的改良，也有研究團隊提出在原架構上嵌入自適應模型或者注意力模塊。例如文獻[10]於 2019 年 MICCAI 比賽中提出嵌合 Selective Kernel Network 的 U-Net 架構，相比其他隊伍之架構並不複雜卻得到與之優異的成績，因此注意力模型或者自適應學習能在高變異度之資料集中得到不錯的成果。

除了辨識與切割的模型架構之外，前處理在醫學影像這種難以取得的小資料集 (small size datasets) 中亦佔一席之地。由於動作偽影 (motional artifacts) 以及能量噪聲 (energy noises) 常出現於心臟 MRI 影像裡，而這些問題會於處理資料正規化時產生不定程度的影像毀損。為消彌這個工程問題，文獻[11]提出以其他方位之心臟切片來定位 ROI (Region of Interest)，同時文獻[12]亦提出以 3D-U-Net 來定位心臟位置同時過濾不重要之訊息與噪聲。此外在擴增資料集方法方面，亦有許多研究團隊以 Generative Adversarial Network (GAN)[13]來合成圖像。GAN 模型可以學習資料分布並以此合成類似影像，這對擁有多種造影模式但數量非常稀少的心臟 MRI 資料集有極大的作用。例如在 2019 年的心臟 LGE 挑戰中，由於主辦單位提供的資料集過小 (小於 50 人) 所以造成各家研究團隊難以訓練以獲得成果。但若用同樣位置不同造影模式之人造 MRI 影像就可以讓深度模型學習到較為貼近真實資料的參數分布，因此合成出的影像會更貼近現實[14]。

有鑑於上述的種種技術發展與呈現，亦由於本計畫所選資料集為小型數據且擁有動作偽影與能量噪聲特性，故將會以定位 ROI 與修改 U-Net 為出發點來進行實驗。

IV. 實驗架構與方法設計

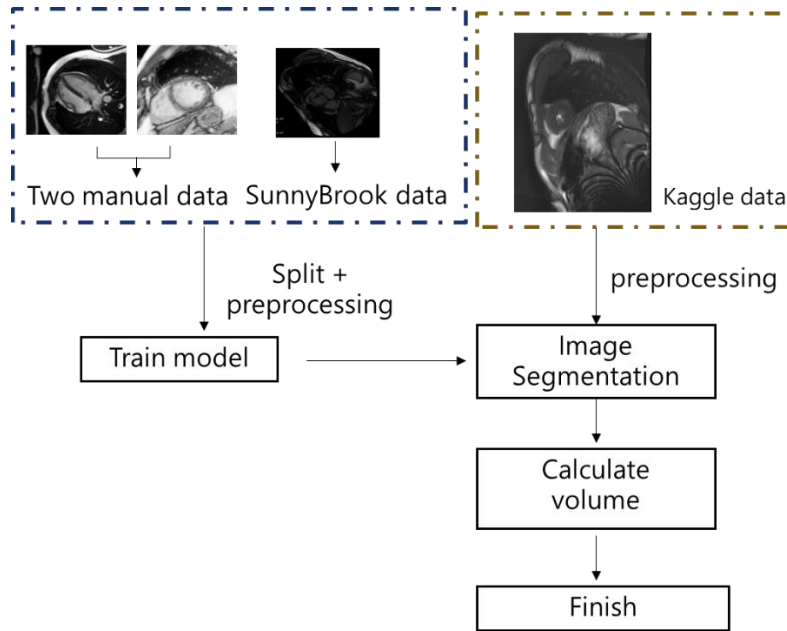
1. 實驗架構：

為構造一個可以自動分割左心室與計算心臟容積的演算法，本計畫先採用 2015 年 Kaggle 左心室定量分析競賽[15]的資料集作為實驗對象。但由於此資料集並未提供 ground truth，因此我們選用當年兩組參賽者[16][17]所標記之資料庫以及更早以前的 Sunny Brook dataset[18]作為訓練集來訓練模型，最後再將模型套用在 Kaggle 資料集並預測左心室容積。本計畫設計之實驗試驗架構如圖一所示。

2. 前處理方法：

✓ 訓練模型之資料集：

將兩個手動標記資料庫與 Sunny Brook dataset 依圖片中心裁切成 176*176 之大小。由於訓練資料集內部差異很大，因此我們亦將 Sunny Brook dataset 的原始



圖一、本次計畫架構圖，訓練資料集(藍色虛線所匡列之圖像)是由三個對比度相差極大的資料集組成，訓練好模型將會用於測試 Kaggle 資料集(棕色虛線之圖像)。

數據使用 histogram equalization 的方式增強圖片的對比度，再將三個資料集經過自適應的方式來調整單張影像的像素強度。自適應的步驟如下：

- (1) 圖像 I 進行最小值最大值正規化(min-max normalization)

$$I_{nor} = (I - I_{min}) / (I_{max} - I_{min})$$

- (2) 正規化後之圖像 I_{nor} 進行 adjust sigmoid function，輸出影像為 I_{out}

$$I_{out} = \frac{1}{1 + e^{mean(I_{nor}) - I_{nor}}}$$

- (3) 由於心臟像素值常大於照片整體平均值，因此我們取 I_{out} 中像素為正的區間再進行最小值最大值正規化，小於或等於零的區間皆為零。

$$I_{out} = \begin{cases} \frac{I_{out} - \min(I_{out})}{\max(I_{out}) - \min(I_{out})}, & \text{if } I_{out} > 0 \\ 0, & \text{if } I_{out} \leq 0 \end{cases}$$

最後再將像素值控制在 0~255 之間。為了評斷訓練模型的優劣，我們將三個資料集平均並隨機打散至訓練資料集、測試資料集與驗證資料集，其比例為 0.8、0.1、0.1，在訓練時也會加入任意旋轉與對比度任意增強之方式來進行資料擴增，驗證後待其模型穩定即會用於 Kaggle 資料集進行預測。

✓ Kaggle 資料集：

Kaggle 資料集分為訓練組 500 人、測試組 440 人，以及驗證組 200 人，此次計畫只選用訓練組的 500 人作為預測對象。而每位病人大致含有以下三種不同方位之圖像：

- (1) 短軸心室圖：估算容積所用之圖像，每人所有之短軸個數不一，每個軸向有 30 張影像序列。
- (2) 兩腔室圖(2 chambers)：每個人有零到一組序列，一組有 30 張。
- (3) 四腔室圖(4 chambers)：每個人有零到一組序列，一組有 30 張。

在資料前處理上由於每位病人有不同大小之圖片尺度與縮放率(pixel spacing)，因此會先定位 ROI 再進行圖像正規化，其方法如下：

- (1) 兩腔室圖與四腔室圖互相垂直之方向即為短軸方向，因此利用此性質可計算出心臟的中心。該病患若無兩腔室與四腔室圖則會以短軸影像中點為心臟中心。
- (2) 為確保在不同縮放率 P 下的影像中心臟皆能在 ROI 內，因此依(1)找出之心臟中心為影像中點(R_{center}, C_{center})上下裁切縮放率為 1.4 公分之區間(D_{row}, D_{col})：

$$D_{row} = R_{center} - 88 * 1.4/P$$

$$D_{col} = C_{center} - 88 * 1.4/P$$

因此新影像上 row 的下界與上界分別為

$$R_{new} = (R_{center} - D_{row}) : (R_{center} + D_{row})$$

而 column 的下界與上界分別為

$$C_{new} = (C_{center} - D_{col}) : (C_{center} + D_{col})$$

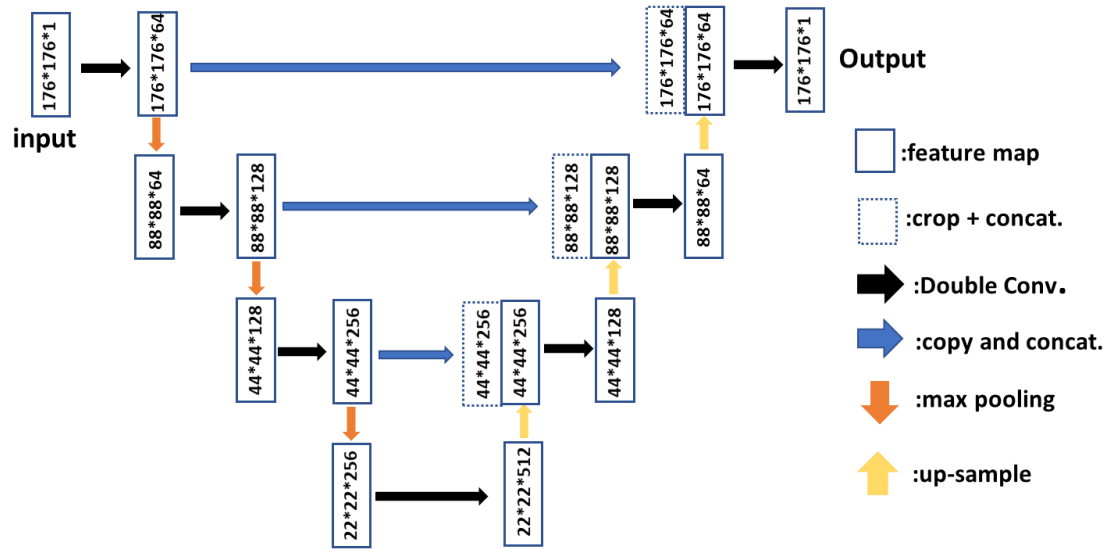
若在裁切時範圍超過影像邊界，則會在邊界外補零再進行裁剪。最後再依此區間之影像縮放為 176*176 之大小，以確保每張影像大小相同，而且影像的縮放率仍為 P。

- (3) 裁剪之圖片經過 histogram equalization 與自適應單元(與訓練模型之資料集的處理方式相同)，最後將像素值控制在 0~255 之間。

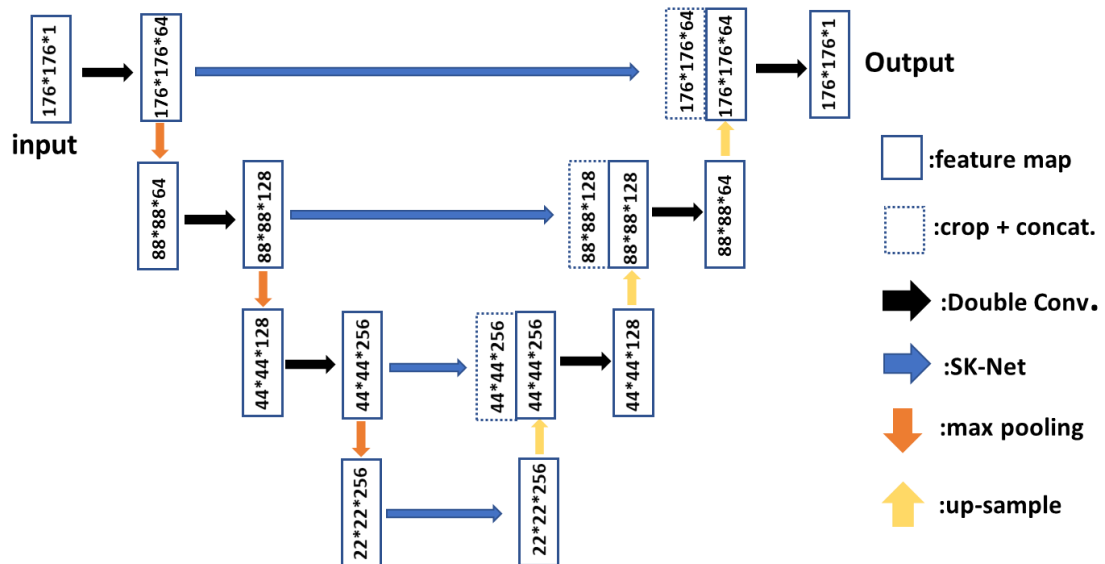
3. 模型架構方法：

本計畫以 U-Net 為主體並嘗試在其中嵌入注意力模型，其中我們加入之注意力模型為 SK-Net 與 CBAM。U-Net 著重在特徵圖(feature map)的空間中擷取特徵，SK-Net 可以找出特徵圖通道中的權重並強化特徵，而 CBAM 則可以在特徵空間與通道單元找出特徵圖之特徵。為了與當代技術做出比較，以下會提供三種測試模型與結果：(1)作為 baseline 的 U-Net、(2)訓練成果較好的 U-Net + SK-Net，以及(3)訓練成果較好的 U-Net + SK-Net + CBAM。此三者之模型架構分別如圖二、圖三與圖四所示。而為公平比較，各個模型中的參數皆相同：batch size 設為 16，loss function 為 tensorflow 內建的 Binary cross entropy，epoch number 設為 60，優化器選用 Adam。而計算模型準確度的標準是語意分割中常用的 Dice 分數，其計算方式如下，其中 X 為 ground truth 而 Y 為預測結果：

$$Dice = \frac{2|X \cap Y|}{|X| + |Y|}$$



圖二、U-Net 架構圖：一般常見的 U-Net 架構是由 5 層的捲積所組成，但由於訓練時發現四層與五層架構正確率並無太大差異，因此本計畫中採用四層的架構。



圖三、U-Net+ SK-Net 架構圖：SK-Net 取代原本 U-Net 的 copy and concatenate 的步驟，可透過學習特徵圖的通道權重來強化語意分割之正確度。

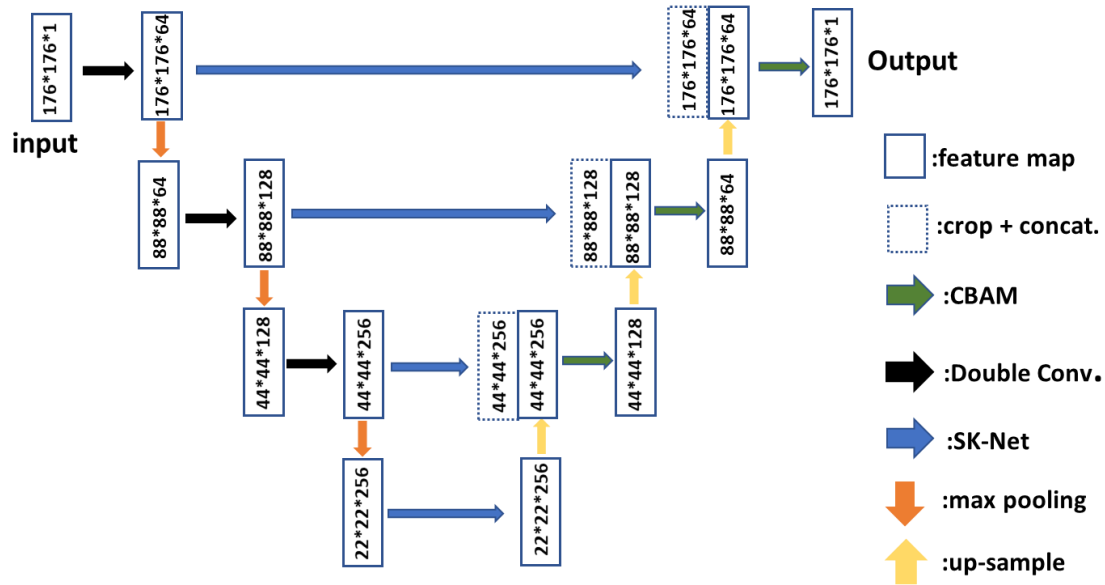
4. 容積估算方法：

將已完成前處理的 Kaggle 資料集放入三種訓練好的模型中預測切割圖，再由切割圖面積 I_{area} 乘上該切片圖像厚度即為單一切片的體積：

$$V_{slice} = I_{area} * I_{thick}$$

而總體積即為該序列下每一切片體積之總合：

$$V_{total} = \sum V_{slice}$$



圖四、U-Net + SK-Net+ CBAM 架構圖：顧慮到 SK-Net 的效果可能在後端解碼層的捲積層所遺失，因此本計畫嘗試在後端加入可在通道與空間上使用的 CBAM 來確保學習權重可傳遞。

而在各短軸序列中體積最大者即為預測舒張體積(diastole)，體積最小者即為預測收縮體積(systole)。最後我們將預測的收縮體積與舒張體積 y_t 與 Kaggle 所提供的正確舒張體積與收縮體積 \hat{y}_t 進行評估，其評斷標準為均方根誤差(root mean square error)：

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{n}}$$

V. 結果與討論

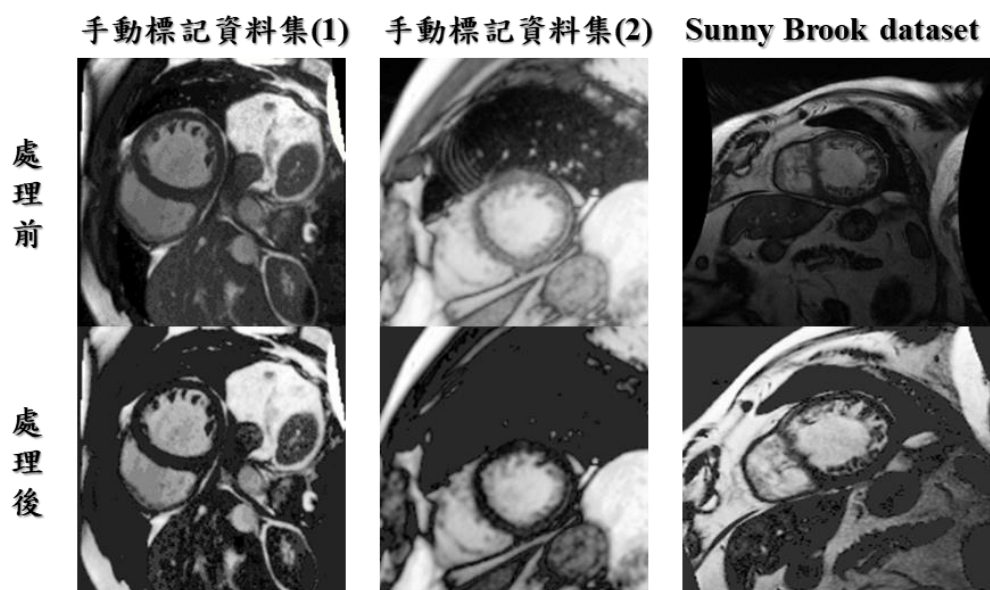
1. 前處理：

訓練資料集前處理之結果如圖五所示，雖然處理前三種資料集의影像對比度差異甚大，但經前處理後可看出此方法除了可以消除對比度差異，也可以完整保留心臟原貌並濾掉其他不重要之臟器。而 Kaggle 資料集(預測資料集)的前處理如圖六所示，從處理結果可看出由多方向的切面(兩腔視圖、四腔視圖)可以大致匡列心臟的位置，而且此方式並不因為高度雜訊干擾而毀損。早期定位 ROI 的方式也使得影像對比度增強時能發揮應有的效果而不受強度高之雜訊干擾。

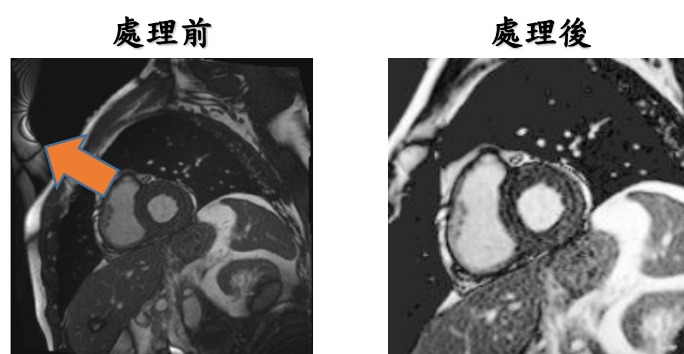
2. 模型訓練結果：

三種模型的 Dice 分數如表一所示，由結果可看出只加入 SK-Net 的模型分數最高，在訓練時也發現加入 CBAM 的模型較不穩定，其訓練的 Dice 有高到如表一所示的 0.94，也有機率掉到比 baseline 的 U-Net 還低的 0.92 左右。

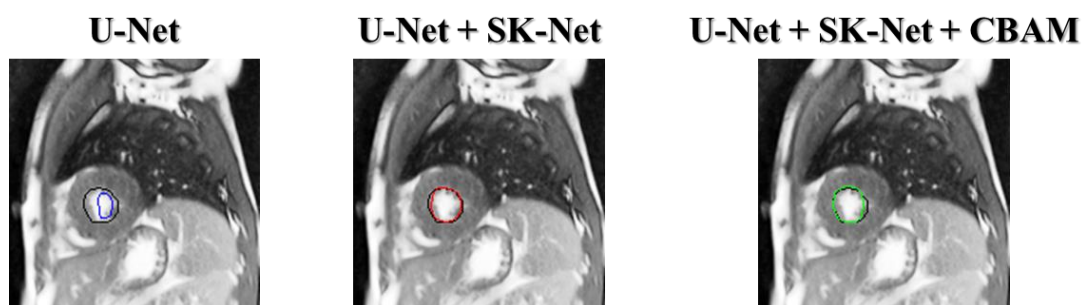
若將上述的三種模型放入測試集中並與手動標記之 ground truth 比較(如圖七所示)，可發現 U-Net 所標記之區域遠小於實際的區域，而加入注意力模型的架構普遍來說可以得到較好之結果。



圖五、訓練資料集前處理之結果，從中可以看出處理前各資料圖像對比度差異甚大，而處理後則趨於一致。



圖六、Kaggle 資料集前處理之結果，處理前影像左上之雜訊在處理後可以完全消失，即可確保像素強度高的雜訊不會影響到後端的自適應對比度校正。



圖七、在訓練資料集中的測試集預測之結果比較圖。資料集中提供的 ground truth 為黑線所匡列之區域，而 U-Net、U-Net+SK-Net 和 U-Net+SK-Net+CBAM 三種模型的預測結果分別為藍色、紅色與綠色所圈選之區域。

表一、三種架構的模型在訓練資料集中各子項資料集的 Dice 分數：

	U-Net	U-Net+SK-Net	U-Net+SK-Net+CBAM
Train	0.9301	0.9552	0.9408
Validate	0.9123	0.9213	0.9205
Test	0.9057	0.9151	0.9148

3. Kaggle 預測結果：

表二為預測 Kaggle 資料集舒張容積與收縮容積之誤差，由表二可以看到三種模型在舒張容積評估中誤差都很大，而在收縮容積評估中誤差雖略小，但也高過比賽中專家建議的標準(RMSE<10ml)。

表二、三種架構的模型在 Kaggle 資料集中預測收縮容積與舒張容積後與正確值的 RMSE：

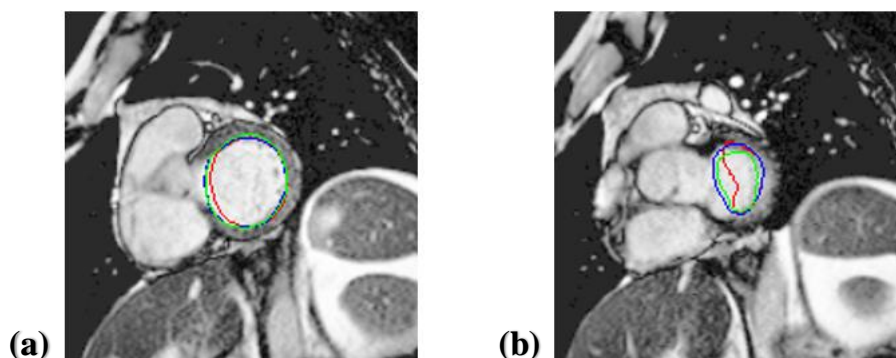
	U-Net	U-Net+SK-Net	U-Net+SK-Net+CBAM
舒張容積誤差(ml)	30.87	31.44	35.91
收縮容積誤差(ml)	15.32	14.36	13.71

4. 討論：以下推論三種模型在 Kaggle 資料集中產生誤差的起因。

(1) 舒張時的預測圖：

圖八為三種模型於 1 號受試者在兩個相同位置但不同序列切片的短軸視圖，而這兩個切片正好位於心臟的中間，因此在總體容積中占很重的比例。圖八(a)是舒張時的切片圖，而圖八(b)是另一個序列的切片圖。由(a)可知基本上在舒張時左心室與其他心房與心室之間之邊界的界線較其他時間點明顯，因此在圖(a)中三種模型預測圖相差不遠；但只要血液灌流到其他腔室(或灌流至心室)時心室與其他腔室的邊界就會開始模糊，此時三者的預測差異就會逐漸拉大。雖然在此切片中圖(b)面積小於圖(a)，但我們並不能保證在其他位置或其他案例中圖(b)的預測面積會小於圖(a)，意旨我們不能保證舒張時所有切片面積皆大於其他時間的切片面積，因為模糊的心臟邊界可能會令模型高估其真實的位置，導致整體舒張容積的上升並使預測容積誤差加大。

在這三種模型中嵌有注意力模塊 SK-Net 與 CBAM 的模型中最容易出現此問題，由於此模型對這類模糊邊界過度敏感，有時甚至會造成不符合人體解剖學的形



圖八、1 號受試者在鄰近舒張時的短軸視圖。U-Net、U-Net+SK-Net 和 U-Net+SK-Net+CBAM 三種模型的預測結果分別為藍色、紅色與綠色所圈選之區域

狀，這或許也可解釋為什麼在訓練時這個模型 Dice 分數的差異會如此之大。而在個體化差異之下我們也發現只要病人的正確舒張體積越大，嵌有注意力模型的舒張容積差異越大，但 U-Net 不論個體差異其誤差都相近。

在 Kaggle 資料集中由於 MRI 的時間解析度(幾秒掃描一張 MRI 影像)與空間解析度(MRI 影像的厚度)在每一個人與位置都不相同，因此每個人舒張容積與正確值的差異也會因這些程不定倍數的增長。由於個體化差異心臟對比度也不同，若我們想要解決此問題，或許更大量的訓練測試集可以解決此問題。

(2) 收縮時的預測圖：

收縮末期時的左心室非常小，有些切面圖中看起來就像一個點或者甚至沒有。而 U-Net 對於此類狀況常常有誤判的狀況導致無法估算切片面積，在心尖與底部的切割預測圖也不甚理想(約 10% 機率無法預測)；而另外兩者由於對邊界敏感度較佳因此在收縮時期預測狀況較為良好，雖有時候心尖或底部切割錯誤，但整體來說與正確體積誤差不大(<5%)。

VI. 結論

在本計畫中我們在前處理中提出了自適應模塊來解決各手動標記資料集中對比度的差異，其在測試集的正確度也證實此方法可以消弭跨資料集中的差距並讓模型學習到圖像特徵。

在模型中我們提出嵌有注意力模型的 U-Net 架構，並以少量手寫標記訓練模型再以大量數據集來做為預測對象。其結果顯示注意力模型對於邊界對比度非常敏感，因此在左心室很小或是心尖與底部等切片中比 U-Net 有更佳的切割成果，但也因其敏感度過高造成高估舒張容積的狀況發生。這項問題可能起因於訓練資料集與 Kaggle 資料集的數量相差甚大，也可能是因為個體化之差異造成心室與其他腔室的邊界對比度相差甚大，若是有更大量的訓練集或許就可解決此問題。

VII. Reference

1. 衛生福利部統計處—歷年死因統計：<https://dep.mohw.gov.tw/dos/np-1776-113.html>
2. Chan, W., Ellims, A. H., Duffy, S. J., Kaye, D. M., & Taylor, A. J. (2012). Principles, current status and clinical implications of ischaemic heart disease assessment by cardiac magnetic resonance imaging. *Internal medicine journal*, 42(1), 7-17.
3. Bernard, O., Lalonde, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P. A., ... & Jodoin, P. M. (2018). Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved?. *IEEE transactions on medical imaging*, 37(11), 2514-2525.
4. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
5. Li, X., Wang, W., Hu, X., & Yang, J. (2019). Selective kernel networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 510-519).

6. Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 3-19).
7. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).
8. Chen, C., Biffi, C., Tarroni, G., Petersen, S., Bai, W., & Rueckert, D. (2019, October). Learning shape priors for robust cardiac MR segmentation from multi-view images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 523-531). Springer, Cham.
9. Khened, M., Kollerathu, V. A., & Krishnamurthi, G. (2019). Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Medical image analysis*, 51, 21-45.
10. Wang, X., Yang, S., Tang, M., Wei, Y., Han, X., He, L., & Zhang, J. (2019, October). SK-Unet: an improved U-Net model with selective kernel for the segmentation of multi-sequence cardiac MR. In *International Workshop on Statistical Atlases and Computational Models of the Heart* (pp. 246-253). Springer, Cham.
11. Luo, G., Dong, S., Wang, K., Zuo, W., Cao, S., & Zhang, H. (2017). Multi-views fusion CNN for left ventricular volumes estimation on cardiac MR images. *IEEE Transactions on Biomedical Engineering*, 65(9), 1924-1934.
12. Xia, Q., Yao, Y., Hu, Z., & Hao, A. (2018, September). Automatic 3D atrial segmentation from GE-MRIs using volumetric fully convolutional networks. In *International Workshop on Statistical Atlases and Computational Models of the Heart* (pp. 211-220). Springer, Cham.
13. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial networks. *arXiv preprint arXiv:1406.2661*.
14. Liu, Y., Wang, W., Wang, K., Ye, C., & Luo, G. (2019, October). An automatic cardiac segmentation framework based on multi-sequence MR image. In *International Workshop on Statistical Atlases and Computational Models of the Heart* (pp. 220-227). Springer, Cham.
15. Kaggle Challenge-Second data science bowl
<https://www.kaggle.com/c/second-annual-data-science-bowl>
16. Kaggle 參賽者手動標記庫(1)
<https://github.com/woshialex/diagnose-heart>
17. Kaggle 參賽者手動標記庫(2)
<https://github.com/woshialex/diagnose-heart>
18. Cardiac Atlas Project-Sunnybrook cardiac dataset
<https://www.cardiacatlas.org/studies/sunnybrook-cardiac-data/>