

# 健康心指南

博儀宸嘉  
盛定佩侑家  
楊高李李郭

# 目錄

1.1 團隊和成員介紹

1.2 項目動機和創意描述

1.3 設計特點和功能

## 1.1 團隊和成員介紹

我們是來自於長庚資管系的"健康心指南"團隊，成員有楊盛博、高定儀、李佩儀、李侑宸及郭家嘉，由於影像醫學的普及，使得醫生判斷病情效率較以前大幅提高，但影像醫學仍有可能會出現難以判讀、甚至醫生彼此見解不一樣的情況發生，這將導致病人及家屬需承受誤診風險，甚至可能有無效醫療的產生。

## 1.2 項目動機和創意描述

現代人由於生活緊湊且壓力大，導致心臟疾病已成為主流文明病。根據衛生署統計，國內大約有 145 萬名心臟疾病患者。65 歲以上老人更每 5 人就有 1 人是心臟病病人，心臟疾病對國人的健康威脅可見一斑。

現代醫療水準日臻完善，醫療儀器的功能也較以往先進許多。但心臟疾病在醫學影像進行標記時，由於疾病位置較無明確定義，因此診斷患者疾病時，常需仰賴醫師們的經驗。然而在現實情況下，並非所有醫師都具有識別心臟超音波圖的資深經驗，要做到準確無誤的判斷，需要醫生數十年以上的經驗，因此判斷病症的高重複性工作，非常適合透過人工智慧訓練來解決專業醫生短缺的問題。而人工智慧另一個優點是縮短了醫生判讀圖像的時間。某些較棘手的疾病中，往往需要數十分鐘才能準確地辨識這些圖像，而準確率也可能因為醫生的疲勞度而下降。然而，有人工智慧的輔助，使模型可以在幾秒鐘內就快速檢測出問題區域，再經由醫生進行最

後確認。提高疾病判斷的準確性和速度，也提高醫療品質。在人工智慧的幫助下，我們期盼未來能夠更準確、更迅速的檢測疾病，提升醫療品質。不僅減少誤診的機率，更減少無效醫療的可能。

本實作透過 Google Colab 平臺進行預測及操作，Google Colab 是一個在雲端執行的編譯環境，由 Google 提供開發者虛擬機，並支援 Python 程式及機器學習演算法。其特點為透過瀏覽器就可以運作，且完全免費。

訓練圖集的醫學影像則使用都普勒超音波圖，都普勒超音波圖特色為利用都卜勒效應判斷血液流速及方向，進而分析臟器功能，但由於超音波圖無完整目標進行界定，因此需要透過模型訓練進行輔助。

最後，模型訓練則採用 Scaled-YOLOv6 架構進行。Yolo 模型為近代知名的神經網路模型，只需要對圖片作一次卷積神經網絡(Convolutional Neural Network, CNN)便能夠判斷圖形內的物件位置與類別，提升辨識效率。在電腦視覺(Computer Vision)領域，目標檢測(Target Detection)一直是一個重要的研究領域，指在一張圖像中找到所有感興趣的目標，其任務包含目標定位和目標分類，同時確定目標的類別和位置，並對特徵進行提取。目標檢測目前廣泛應用於人臉識別、自動駕駛、智慧交通、車牌識別、醫學影像等領域，具有十分重要的研究意義。

經由模型訓練及預測，並在上傳圖片或影片後進行判讀。本實作目標為利用深度學習(Deep Learning) 架構輔助醫師有效率地進行診斷，且同時

提供網站的客製化功能開放給大眾使用。

本實作替兩類常見的心臟疾病：心室中膈缺損(Atrial Septal Defect, ASD)與心房中膈缺損(Ventricular Septal Defect, VSD)，提供 AI 輔助檢測系統。並架設網站提供給大眾使用。

### 1.3 設計特點和功能

#### YOLO 介紹

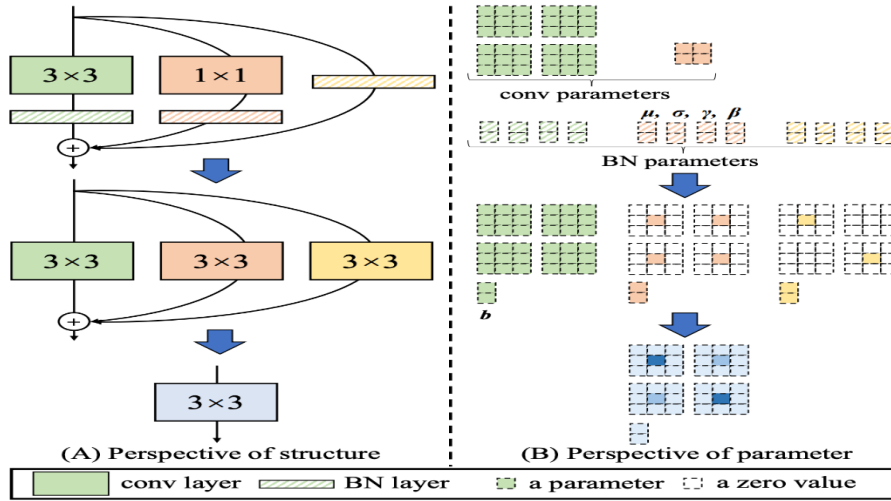
YOLO 的全名是 You Only Look Once，其名字顧名思義就是你只需要做一次卷積神經網路 (Convolutional Neural Network, CNN) 就可以辨別出圖片中物體的類別和位置。其主打特點為即時檢測(real time detector)。一般來說，依靠電腦辨識圖片需要做到以下兩大區塊:單一物件處理以及多重物件處理，而這兩大區塊又可細分為分類(Classification)、定位(Localization)、物件辨識(Object Detection)及實體分割(Instance Segmentation)，但是其缺點為耗時且耗資源，或為了讓檢測更精準，前置的資料標記作業更為繁瑣。

過往的物件辨識採用兩階段式(two stage)分工，也就是使用兩個神經網路，第一個神經網路負責找出圖片中的目標物件，並進行標記後再給下一個模型進行分類。YOLO 則採用單階段式方法，單階段式(one stage)方法是指只使用一個神經網路提取一次特徵即可實現目標檢測，大大提升辨識速度。再來我們將會對 YOLOv6 及 YOLOv7 做介紹。

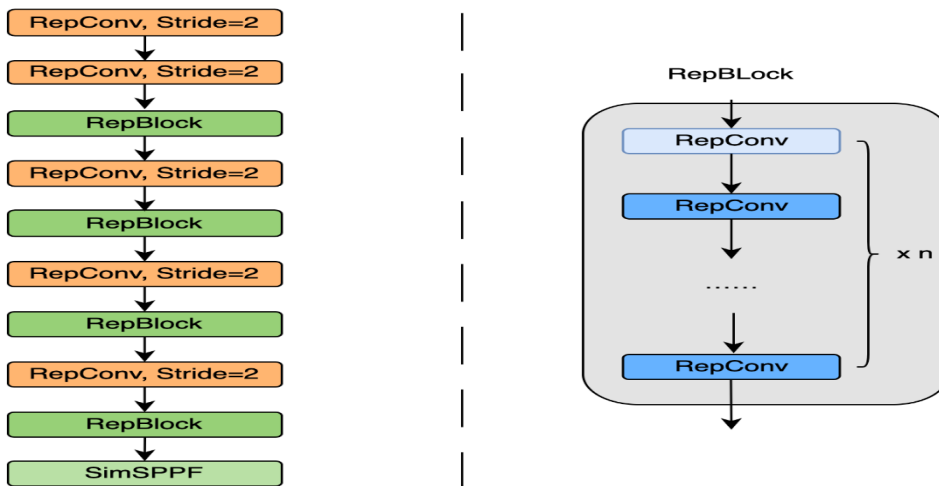
## YOLOv6 介紹

YOLOv6 與其他 YOLO 系列最大不同在於：YOLOv6 改善了模型的訓練策略以及其網路架構。另外在部署方面 YOLOv6 支持多個不同平臺進行部署，極大地簡化操作流程及配置工作。

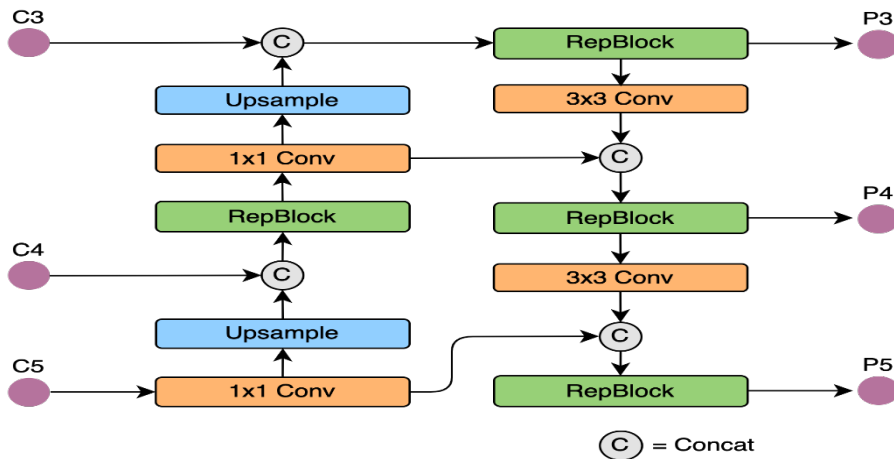
以 Backbone 及 Neck 來說，YOLOv5 採用多分支的殘差結構 CSPNet[1] 搭建。但此技術對於某些硬體來說，會增加一定程度地延遲，同時減少記憶體使用率。針對此方面的問題，在 YOLOv6 方面，引用了可重參數化、更高效的骨幹網路 EfficientRep 及 Rep-PAN Neck，其特點在於：引用 RepVGG[2] 結構以及重新設計 Backbone 和 Neck。RepVGG 結構是一種在模型訓練時具有多個分支拓撲，另外在實際配置時可融合成 3x3 卷積的可重參數化結構（如圖一）。通過上述策略，YOLOv6 可大幅減少硬體的延遲，同時也提升了精確度；接著 Backbone 中 stride=2 的捲積層更換成 stride=2 的 RepConv 層，同時將 CSP-Block 重新設計為 RepBlock(如圖二)；最後在 Rep-PAN Neck 設計方面，為了達到更好的精度與速度平衡，YOLOv6 以 RepBlock 取代 YOLOv5 的 CSP-Block，同時調整 Neck 運算，使硬體進行有效推理同時，維持較好的多尺度特徵融合能力(如圖三)。



圖一、RepVGG 融合過程



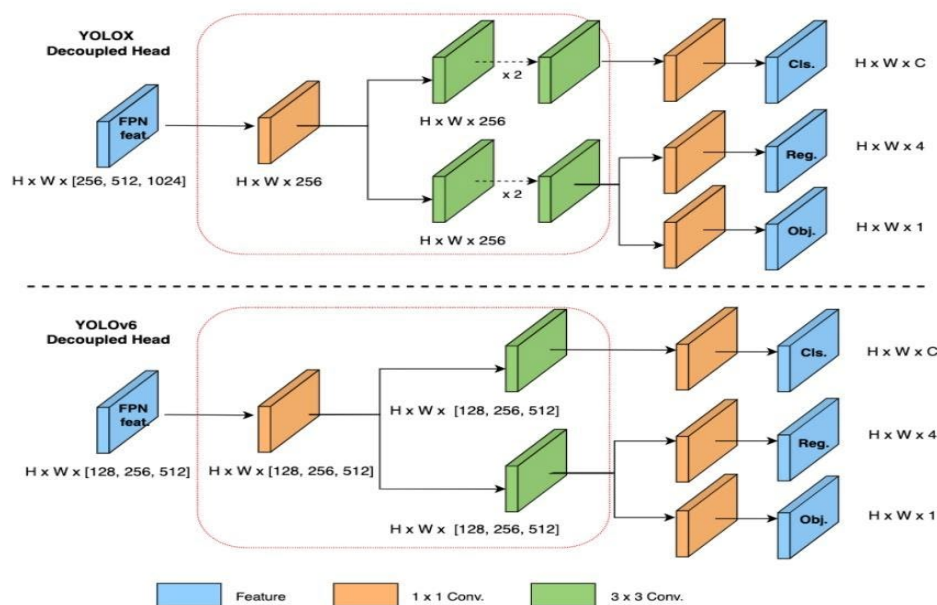
圖二、EfficientRep Backbone 結構圖



圖三、Rep-PAN 結構圖

在 Head 部分，YOLOv6 設計出比 YOLOv5 更簡潔有效的解耦偵測頭

(Decoupled Head)，並對其進行精簡設計。YOLOv6 採用 Hybrid Channels 策略重新設計出一個更高效的解耦頭結構。在維持精確度同時，也可以降低一般解耦頭所帶來的額外時間成本，達到速度與精度的權衡。



圖四、Decoupled Head 結構圖

## YOLO 訓練模式

在訓練方面，YOLOv6 採用更簡便的 Anchor-free 檢測方法。過去進行目標檢測大多是基於 Anchor，但使用 Anchor 會出現以下幾點問題：

(1)訓練前需要進行集群分析(Cluster analysis)，找出最佳 Anchor 集合，但這些集合的資料集彼此相關，因此泛化 (Generalization) 性能較差。

(2)使用 Anchor 會提高偵測頭的複雜度。

相較於 Anchor，Anchor-free 的泛化能力更強，且解碼邏輯簡單，另外 Anchor-free 偵測器在速度上較 Anchor 有 51%的提升。



此外，為了取得更多高質量正樣本，YOLOv6 引入 SimOTA，為不同目標設定不同的正樣本數量，同時也進一步提高偵測精確度。因為在偵測模型推理過程中難免會出現正負樣本不平衡的現象，所以我們需要對正負樣本進行平均。

為了更進一步提升迴歸精度，YOLOv6 採用 SIOU[3]邊界框(Bounding Box)迴歸損失函數來監督網路學習。目標檢測網路訓練需要至少定義兩個損失函數：邊界框迴歸損失以及分類損失，而損失函數的定義往往會對檢測及訓練速度產生較大的影響。

## YOLOv7 介紹

為了設計出最佳化的模組與方法以提升物件偵測的準確度，同時降低推論的成本(inference cost)以實現即時物件偵測器能夠同時支持邊緣及雲端的行動 GPU 與 GPU 裝置，在 YOLOv6 上架兩星期後，YOLOv4 團隊更新了模型架構，並命名為 YOLOv7。

與 YOLOv6 及其他當前主流的即時目標檢測器不同，YOLOv7 考慮到「如何設計一個高效網絡」的問題。Wang et al.在模型架構方面提出一個延伸 ELAN(Extended-ELAN, E-ELAN)架構，使其能夠在不破壞原本的梯度路徑(gradient path)的情況下不斷提升網路的學習能力，此外 E-ELAN 僅利用組卷積放大計算塊(computational block)的通道及基數(cardinality)，對於過度層(transition layer)則未做更改，使得 E-ELAN 不僅可以維持原本

ELAN 的設計架構，即藉由控制最短與最長的梯度路徑令深度網路可以有效率地學習與收斂(converge)，也可引導計算塊的不同組學習更多元的特徵；而在針對序連模型(concatenation-based model)的模型縮放方面提出一個複合縮放方法(compound scaling method)以解決序連模型在進行深度的縮放(depth scaling)時，其輸出的計算塊的寬度(width) 以及接續的轉換層(transmission layer)的輸入寬度會增加的問題，而透過實驗表明複合縮放方法比起單一縮放寬度可以提升 0.5\% 平均精準度(AP)，也可減少參數的使用與計算量，但在比起單一縮放深度方面則需額外多使用參數 2.9\% 與計算量 1.2\% 才可獲得 0.2\% AP 的提升；除此之外在利用梯度流傳播路徑(gradient flow propagation paths)分析後發現 RepConv 中的 id 連結(identity connection)會破壞 ResNet 的殘差(residual)與 DenseNet 的序連以至於為不同的特徵圖(feature maps)提供更多元的梯度的問題，因此提出計畫再參數模型(planned re-parameterized model)移除 id 連結，即由 RepConvN 取代 RepConv，而透過計畫再參數模型在與殘差模型(residual-based model)比較實驗發現兩者並無差異；另外在深度監督標籤分派(label assignment)方面有別於分別利用協助訓練的頭(auxiliary head, 輔助頭)與最終輸出的頭(lead head, 領頭)的自身預測結果以及基準真相(ground truth)去執行分派，其提出一個新的深度監督標籤分派方法，包括領頭引導標籤分派器(lead head guided label assigner)與從粗到細的領頭引導標籤分派器(coarse-to-fine lead

head guided label assigner)，這些分派器可以被領頭的預測結果與基準真相 (ground truth) 進行最佳化以同時取得訓練時的領頭與輔助頭的標籤，而實驗上也發現領頭引導標籤分派在 AP、AP50 與 AP75 上均比起普通獨立標籤分派 (general independent label assignment) 來的高，分別為 55.9%、73.5% 與 61.0%，另外在從粗到細的領頭引導標籤分派上則有最好的表現，在 AP、AP50 與 AP75 的值分別為 55.9%、73.5% 與 61.1%。

在物件偵測實驗方面，YOLOv7 與 YOLOv4 比較之下，YOLOv7 的參數減少 75%，而計算量也減少 36%，但是 AP 卻提升 1.5%，若與 YOLOR-CSP 相比，YOLOv7 的參數減少 43%，計算量降低 15%，提升 0.4% AP，若與微模型 (tiny model) YOLOv4-tiny-31 比較，YOLOv7-tiny 的參數減少 39%，且計算量也少 49%，但 AP 並無改變；另外在與各自不同的最先進的目標檢測器 (state-of-the-art object detectors) 對比方面，不管在幀率 (FPS)、AP、參數使用率與推論速度上，各調整後的 YOLOv7 於表現均有程度上的提升。另外由於 YOLO 官方無釋出 YOLOv6 及 YOLOv7 的比較，故不加以贅述。

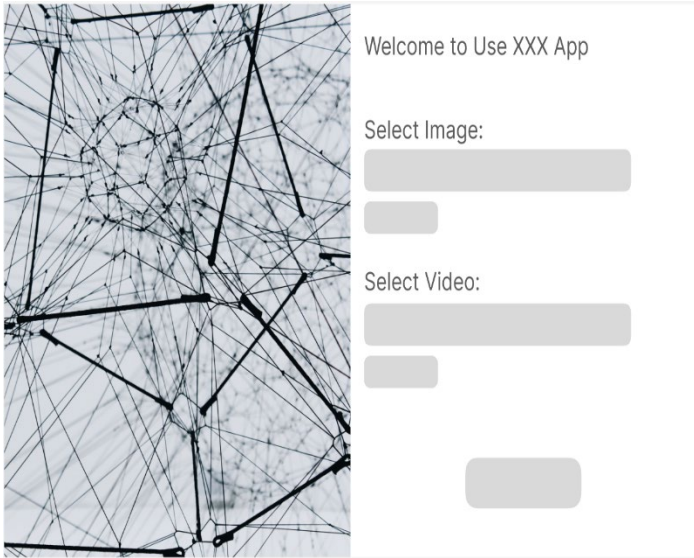
Model	#Param.	FLOPs	Size	AP <sup>val</sup>	AP <sup>val</sup> <sub>50</sub>	AP <sup>val</sup> <sub>75</sub>	AP <sup>val</sup> <sub>S</sub>	AP <sup>val</sup> <sub>M</sub>	AP <sup>val</sup> <sub>L</sub>
<b>YOLOv4</b>	64.4M	142.8G	640	49.7%	68.2%	54.3%	32.9%	54.8%	63.7%
<b>YOLOv4-CSP</b>	52.9M	120.4G	640	50.3%	68.6%	54.9%	34.2%	55.6%	65.1%
<b>YOLOv7</b>	36.9M	104.7G	640	<b>51.2%</b>	<b>69.7%</b>	<b>55.5%</b>	<b>35.2%</b>	<b>56.0%</b>	<b>66.7%</b>
improvement	-43%	-15%	-	+0.4	+0.2	+0.2	+1.5	=	+1.3
<b>YOLOv4-tiny</b>	6.1	6.9	416	24.9%	42.1%	25.7%	8.7%	28.4%	39.2%
<b>YOLOv7-tiny</b>	6.2	5.8	416	<b>35.2%</b>	<b>52.8%</b>	<b>37.3%</b>	<b>15.7%</b>	<b>38.0%</b>	<b>53.4%</b>
improvement	+2%	-19%	-	+10.3	+10.7	+11.6	+7.0	+9.6	+14.2
<b>YOLOv4-tiny-3l</b>	8.7	5.2	320	30.8%	47.3%	32.2%	<b>10.9%</b>	31.9%	51.5%
<b>YOLOv7-tiny</b>	6.2	3.5	320	<b>30.8%</b>	<b>47.3%</b>	<b>32.2%</b>	10.0%	<b>31.9%</b>	<b>52.2%</b>
improvement	-39%	-49%	-	=	=	=	-0.9	=	+0.7
<b>YOLOv4-E6</b>	115.8M	683.2G	1280	55.7%	73.2%	60.7%	40.1%	<b>60.4%</b>	69.2%
<b>YOLOv7-E6</b>	97.2M	515.2G	1280	<b>55.9%</b>	<b>73.5%</b>	<b>61.1%</b>	<b>40.6%</b>	60.3%	<b>70.0%</b>
improvement	-19%	-33%	-	+0.2	+0.3	+0.4	+0.5	-0.1	+0.8
<b>YOLOv4-D6</b>	151.7M	935.6G	1280	56.1%	73.9%	61.2%	<b>42.4%</b>	60.5%	69.9%
<b>YOLOv7-D6</b>	154.7M	806.8G	1280	56.3%	73.8%	61.4%	41.3%	60.6%	70.1%
<b>YOLOv7-E6E</b>	151.7M	843.2G	1280	<b>56.8%</b>	<b>74.4%</b>	<b>62.1%</b>	40.8%	<b>62.1%</b>	<b>70.6%</b>
improvement	=	-11%	-	+0.7	+0.5	+0.9	-1.6	+1.6	+0.7

圖五、YOLOv7 各尺寸模型與其他模型之比較

## 網站操作流程

我們利用 Figma 進行網站架設，由於專題是針對心臟逆流圖像暨影片進行預測，因此對於介面設計的部分可分成兩個畫面(圖六及圖七)。

圖六的網站功能為可令使用者選擇上傳圖片或影像以進行預測。主要針對以預測好並且產出的圖片進行呈現。而圖七對於預測好所產生的圖片，可於圖中大框框呈現。「返回」鈕則可讓使用者返回圖六，無需再重新啟動應用程式。



圖六、網站示意圖



圖七、網站示意圖

#### 1.4 綠色理念

針對人工智慧輔助診斷，不但能夠減少誤診率、大幅減少診斷時間之外，最重要的是能夠減少醫療浪費，達到環保理念。

參考文獻：

- [1] CSPNet : A New Backbone that can Enhance Learning Capability of CNN, [https :  
//arxiv.org/abs/1911.11929](https://arxiv.org/abs/1911.11929)
- [2] RepVGG : Making VGG-style ConvNets Great Again, [https :  
//arxiv.org/pdf/2101.03697](https://arxiv.org/pdf/2101.03697)
- [3] SIoU Loss : More Powerful Learning for Bounding Box Regression, [https :  
//arxiv.org/abs/2205.12740](https://arxiv.org/abs/2205.12740)