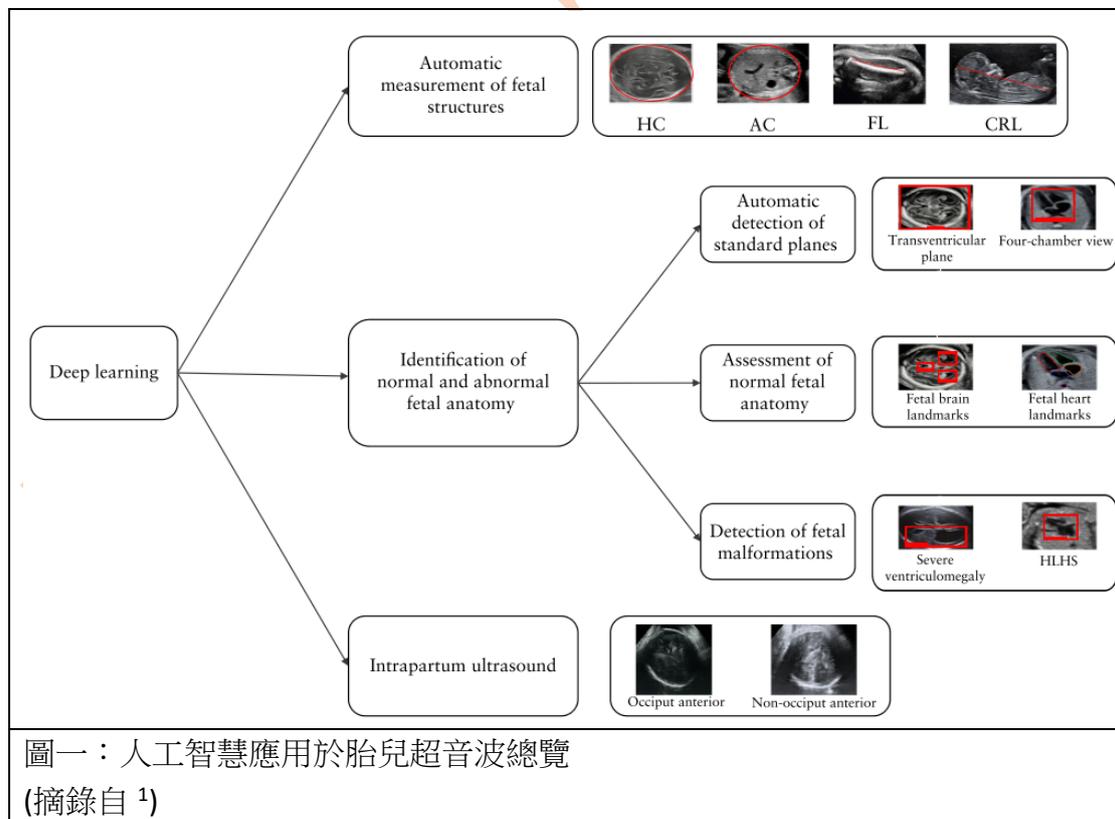


一. 簡介

近幾年由於電腦運算能力進步的原因，人工智慧 (Artificial intelligence, AI) 變得相當熱門，相當多的 AI 應用及研究如雨後春筍般地出現，其中也包含了應用於胎兒超音波的人工智慧演算法。在 ISUOG 的一篇文獻中¹有統整了近期人工智慧應用於胎兒超音波的應用，分為下列三大類(圖一):

1. 自動測量胎兒構造，如自動測量胎兒頭圍 HC、腹圍 AC、股骨長 FL、及頭臀長 CRL
2. 辨認正常與不正常之胎兒構造，如自動辨認標準切面 (如找出頭部的 Transventricular plane)、評估正常胎兒的構造 (如心臟四腔影像中各個心房心室的特徵)、辨認異常的構造 (如頭部嚴重腦室擴大、心臟左心發育不全症候群)
3. 產中超音波 (Intrapartum ultrasound)，例如自動辨別是否為 OA 胎位



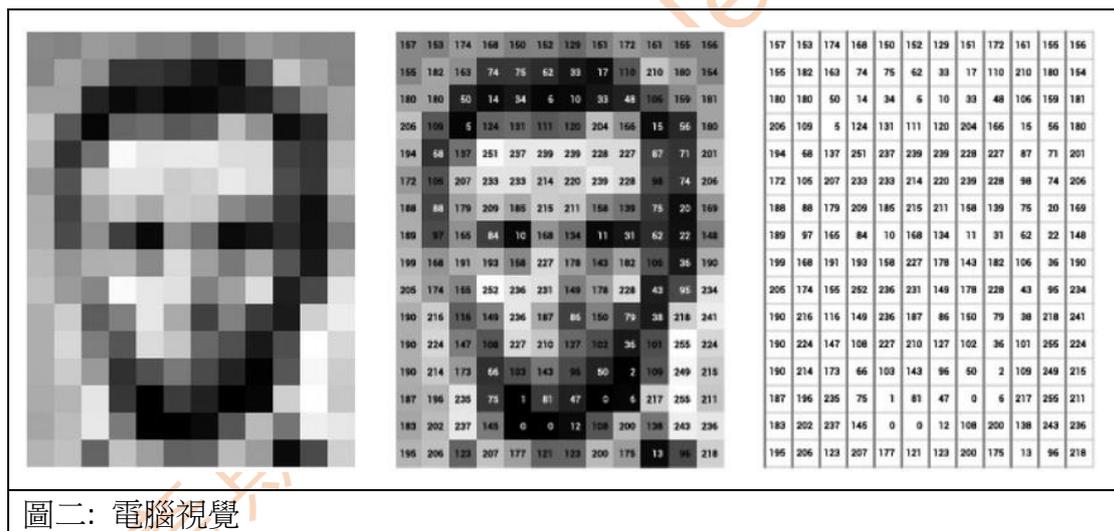
圖一：人工智慧應用於胎兒超音波總覽 (摘錄自¹)

二. 電腦視覺 (Computer vision)

電腦對於圖像的判讀和人腦相當不一樣，如下圖二所示，左邊的圖是一張黑白的林肯像，我們可以很直觀地判讀這個圖像是什麼，機器則是判讀成右邊的二維矩陣 (12x15 共 180 個數字)，這二者有什麼關係？

由電腦判讀的數位圖像，均以像素 (Pixel) 作為基本的單位，數值介於 0 到 255，共 256 個數值 (8 bit 色深)，其中 0 是最黑，255 是最白，1-254 便依照數字大小從黑到白遞增，如中間的林肯像，便是以色深和數字去做對照，因此一張 12x15 像素的影像，對電腦來說便是 12x15 的二維矩陣。

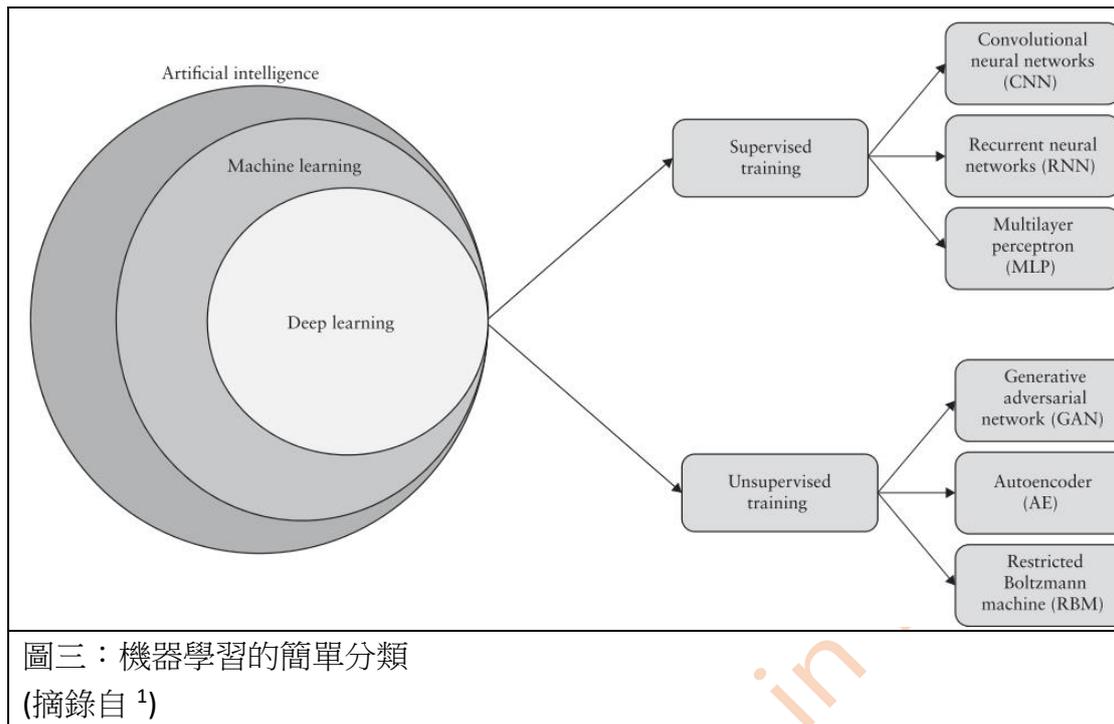
黑色影像只有黑白程度的差異，而彩色影像則是由 RGB (Red, Green, Blue) 三原色所組成，因此有三個通道 (Channel)，若是把林肯像變成彩色影像，則會有三個矩陣二維矩陣，意即紅色一個、綠色一個、藍色一個，因此彩色影像會從二維矩陣變成三維矩陣；舉例來說，一張 Full HD (長 1920 像素、寬 1080 像素) 的彩色影像，對電腦來說便是一個 1920x1080x3 的矩陣，共有 6,220,800 個數字介於 0~255 的三維矩陣。



圖二：電腦視覺

三. 人工智慧基本知識

人工智慧 (AI) 是一種機器學習演算法的範疇，它可以自動地去做特徵提取 (feature extraction)，並從學習的資料來改進演算法，進而提升模型的準確率，圖二即為機器學習的概觀：



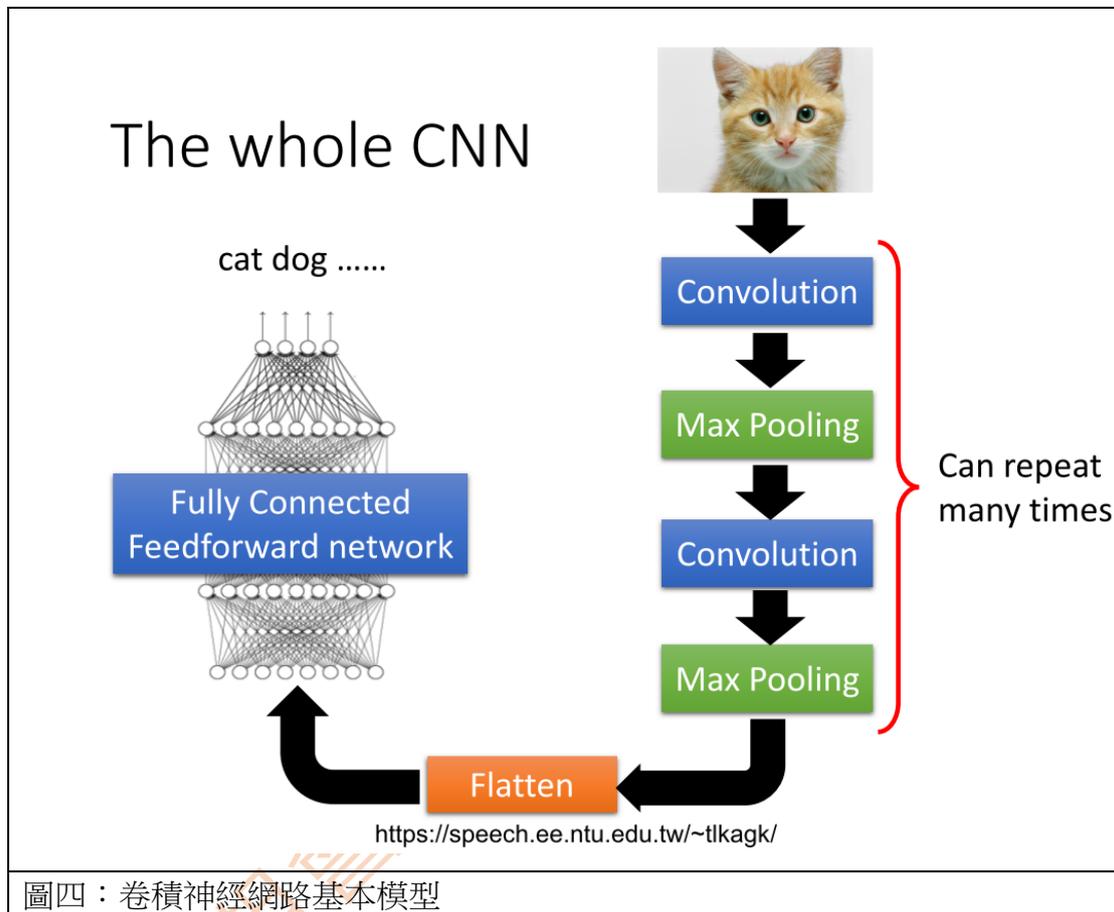
目前在胎兒超音波中，應用最多的機器學習法為深度學習，而它包含了二大類：監督式學習 (Supervised learning)，及非監督式學習 (Unsupervised training)。非監督式學習指的是讓機器自己依學習資料的特徵去做分群 (Clusters)，例如若是要區分狗和雞這二種動物，機器會去分為二隻腳的一群 (雞)，四隻腳的一群(狗)，不用經過人類標記；反過來說，監督式學習需要由人類來標記出位置在哪及正確答案 (Ground truth) 是什麼，例如若是要辨識道路上的各種物件，需要由人類標記出汽車、機車、行人、紅綠燈等物件的位置，再由機器學習後後才能去判讀，無法不經標記而得到結果。

在胎兒超音波中，通常我們會需要判斷正常與不正常的切面與構造，更進一步的是要去看構造位在圖像中的哪裡，例如是否有看到透明中隔(CSP)？透明中隔的位置又在哪裡？在這種情況下，大部分應用於胎兒超音波深度學習為監督式學習 (需要人為標記)，其中又以卷積神經網路 (Convolutional Neural Network, CNN) 最被廣泛應用。

卷積神經網路 (CNN) 的架構如圖四，若我們是要訓練一個可以判讀貓圖像的神經網路，首先會把貓圖像(對電腦來說是三維數字矩陣)輸入神經網路，此圖像會先進入卷積層 (Convolution layer, CONV)，卷積層裡有設計好的濾波器 (kernel) 來擷取影像的特徵，擷取後的特徵接著會進入池化層 (Pooling，最常用的是最大池化 Max Pooling, MP)，這層的目的是要保留特徵資料的同時，也減少資料的大小來減少運算量，其中卷積層 (CONV) 和最大池化層 (MP) 可以合在一起重複很多次 (深度學習，深度即很多層、很多次的意思)；在經過多次的卷

積及池化後，擷取特徵後的資料量已經比原本圖像 (數字矩陣) 小很多，於是可以進入全連接層 (Fully connected feedforward network, FC) 來做推理連接，以判別圖像是不是貓；上述的步驟總結來說，卷積神經網路(CNN)的架構可簡化為：

圖像輸入→(CONV+MP)*n→ Flatten → FC → 判讀結果

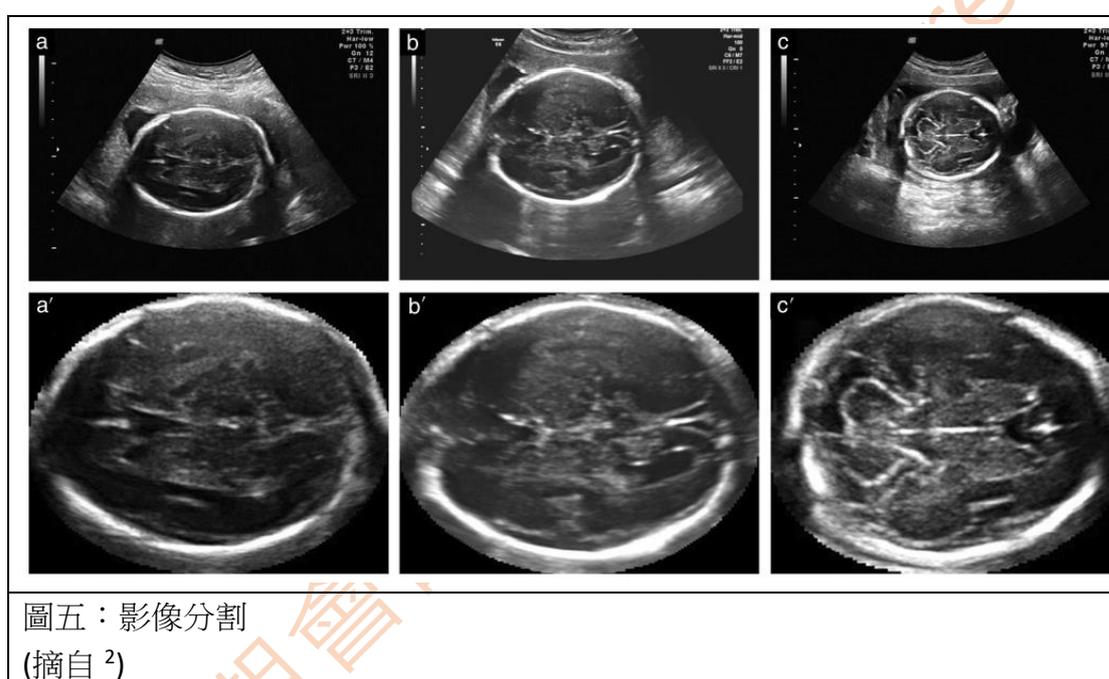


目前已經有許多前人訓練好的圖像判讀模型(Pre-trained model)，如 AlexNet、VGG19、GoogLeNet、YOLO 等模型，這些模型在之前的圖像辨別競賽中脫穎而出，代表其架構優秀，因此可以用來做轉移式學習(Transfer learning)；轉移式學習是將已經訓練好的神經網路模型，保留其主要架構，再利用新的圖像來做訓練，像是利用超音波影像，訓練既有的模型做新的分類及判讀。

四. 人工智慧之應用

人工智慧用於胎兒超音波影像，可以簡單地分做四種功用，分類 (Classification)、定位(Localization)、物件偵測(Object detection)、影像分割 (Segmentation)，簡述如下：

1. 分類:將圖像分為正常或不正常，或是分類出是哪一個切面(如 Four-chamber view 還是 LVOT)，通常不知道機器的判讀依據是什麼。
2. 定位: 找到圖像上的構造，如找到腦部切面圖像中的 CSP、側腦室在哪裡
3. 物件偵測: 分類及定位的總合，例如對於胎兒心臟影像，首先機器要找出心臟的各個構造(心室、心房、肺靜脈等，此步驟為定位)，再辨識出該影像為心臟切面(分類)，
4. 影像分割：機器找出影像中重要的部份，將其保留，刪去影像中其它不重要的部分，例如在胎兒頭部影像中，只保留畫面中頭部的影像，將頭外面的羊水、胎盤等構造消除，如下圖五



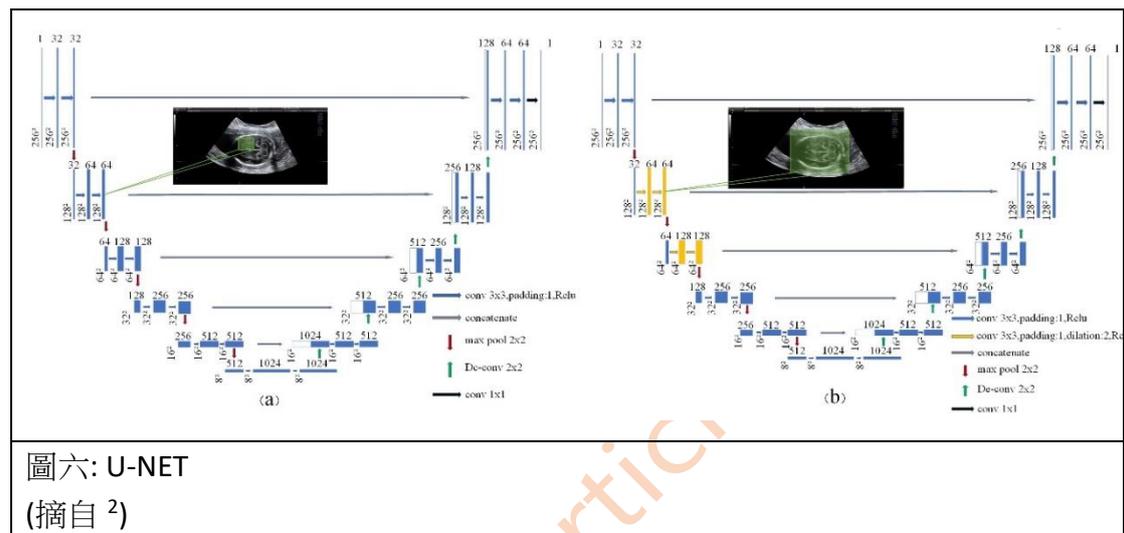
五. 文獻分享

目前已有許多人工智慧與胎兒超音波相關的研究，本文使用一篇發表在 ISUOG 上的文獻做為範例來簡介 AI 是如何應用於胎兒超音波。

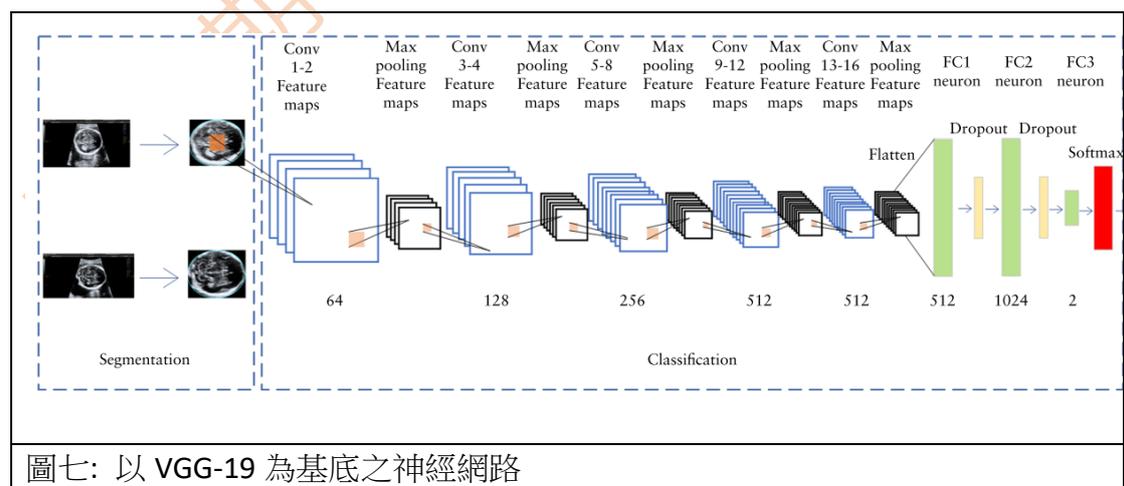
這篇發表在 ISUOG 的文獻(2020)²，Xie 等人使用單一醫學中心的影像資料來訓練機器判讀胎兒頭部的超音波標準切面，分為正常與不正常(分九種，包含 Neural tube defect, holoprosencephaly, midline structural anomaly, lissencephaly, microcephalus, posterior fossa anomaly, space-occupying lesion, intracranial hemorrhage 及 ventriculomegaly)，共 10 種分類。作者納入懷孕週數 18-32 週的個案，由 15 位超音波專家收案，共納入了 92,784 個案，在扣除不符合收案標準的個案後，共有 15,372 正常及 14,047 異常個案納入，資料集以 8:2 的比例分

為訓練集及測試集(在訓練模型時，通常會用八成的資料來訓練、優化模型，再以二成資料來驗證模型成效)；其中，判斷正常或不正常的依據為新生兒的超音波，或者是屍體解剖(此即 **Ground truth**)。

作者先使用 **U-Net** 模型(常用來做影像分割的模型，架構像是一個 **U** 字型而得名，如圖六)對超音波影像進行影像分割(**Segmentation**)，如上圖五，將影像分割出只剩頭部區域，減少雜訊，再輸入以 **VGG-19** 為基底的卷積神經網路(**CNN**)做轉移式學習與訓練(圖七)。

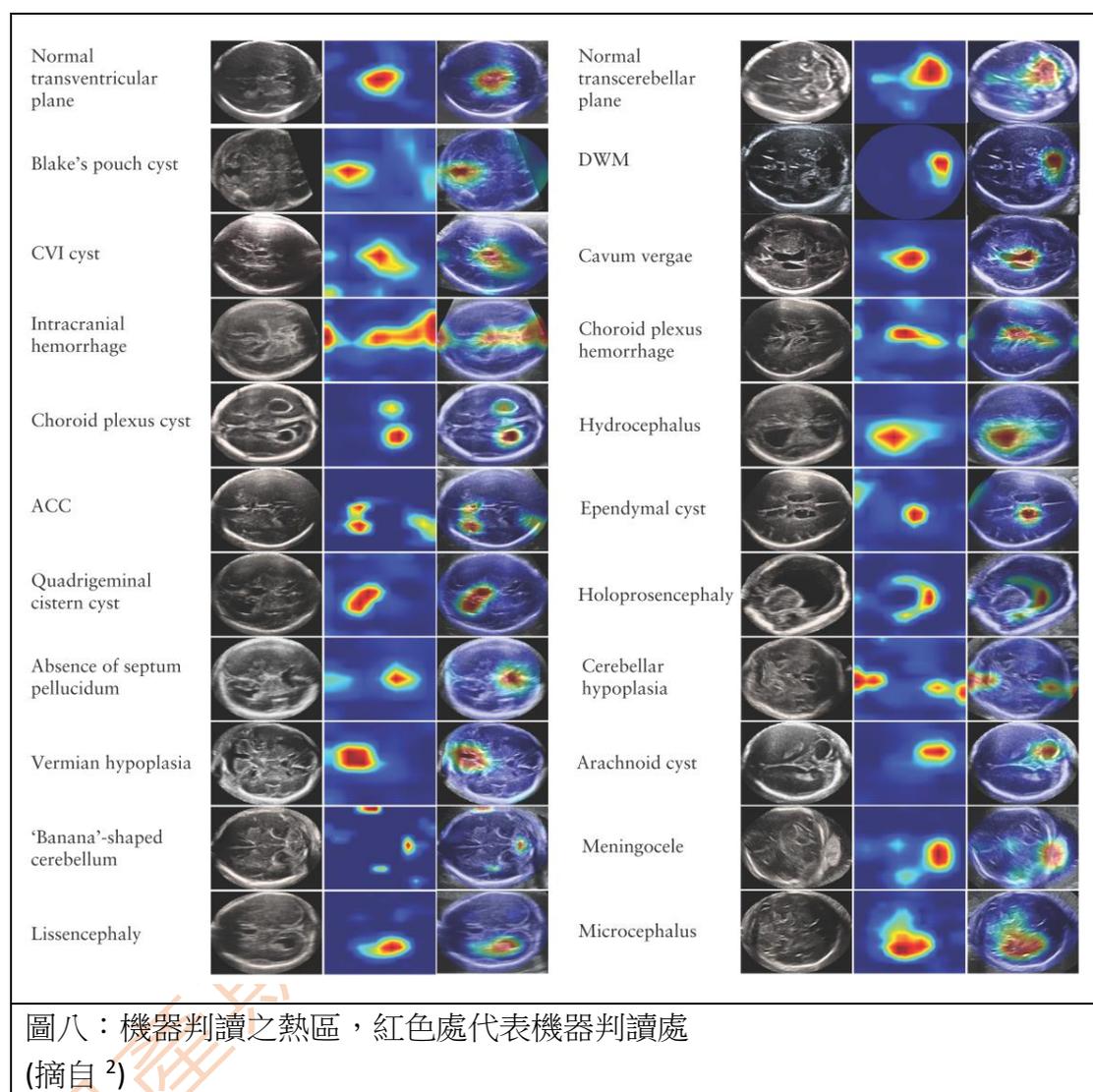


經過分割後的影像會輸入進神經網路，在卷積層(**CONV**)及最大池化層(**MP**)共 **16** 層的處理後，資料量已變小許多，因此可將此圖像(二維數字矩陣)攤平 (**Flatten**)成一個直向量，再經過三層的全連接層去計算，最後算出該影像屬於哪個分類(正常或是九種異常中的哪一種)，過程如圖七所示。



人工智慧有個難解的問題，在於我們不知道機器是如何去運算而得出這個結果，因而會把機器學習的過程稱為黑盒子(**Black box**)，意即看不出它是怎麼算

出答案的，如同一個數學題目，只知道機器算出的答案正確，卻不知算式為何一樣。因此有學者開發出一個演算法，**Grad-CAM(Gradient-weighted Class Activation Mapping)**，透過 CNN 在計算時數字梯度的改變，來看機器是在哪裡判讀出問題，並以熱區表示，如下圖八。



經過訓練之後，此模型判讀的能力相不錯，對於正常的個案正確率有 95.7%，而異常個案的正確率為 96.9%，AUC 有 0.989；有趣的是，即便正確率高，使用熱區來看機器是否判讀在正確的位置(例如胼脂體發育不全之個案，機器應該在 CSP、Lateral ventricle 等地方為顯示為判讀熱區，不應在其它像是小腦或頭骨處顯示為判讀熱區)，在人類專家評估機器覺得異常的地方，只有 86.2% 是落在正確位置，13.7%則是落在完全不相干的位置，表現並不是太好。

六. 總結

目前有愈來愈多優秀的人工智慧模型橫空出世，例如生成式 AI 的 ChatGPT 等，未來使用這些模型的應用將會愈來愈精彩，例如或許可以使用卷積神經網路 CNN 判讀圖片，再由物件偵測模型 YOLO 找出構造位置，經由 ChatGPT 類型的生成式 AI 打報告，最後由人類來檢視結果，此模式將會大幅減少臨床作業的時間，也可減少常規作業所造成的疲勞，讓我們有更多時間去做更具創造力的事情。

由於目前人工智慧應用於超音波仍需人類操作超音波來餵影像給機器去判讀，另外，一些比較困難的先天性異常診斷(例如區分都有 VSD + Over-riding aorta 的 CHD，是 TOF 還是 Truncus arteriosus)，異常個案的諮詢(是 isolated 異常，或者是 genetic syndrome)，仍需人類專家的經驗，因此在可預見的未來，人工智慧無法取代我們，但人工智慧可以扮演一個很好的輔助角色。

Reference:

1. Ramirez Zegarra R, Ghi T. Use of artificial intelligence and deep learning in fetal ultrasound imaging. *Ultrasound Obstet Gynecol* 2023;62:185-94.
2. Xie HN, Wang N, He M, et al. Using deep-learning algorithms to classify fetal brain ultrasound images as normal or abnormal. *Ultrasound in Obstetrics & Gynecology* 2020;56:579-87.