摘要：

這篇論文使用類神經網路來計算胎兒頭圍指標(HC)。

在胎兒超音波影像裡，有各種評估胎兒健康狀況的指標，其中最明顯且好估計的指標是HC，我們藉由類神經網路的影像分割模型中的Unet3+並結合Attention Gate來達到精準的分割，得到分割後再由橢圓擬合的最小二乘法算出最後HC的輸出。

介紹：

超音波造影比起他造影方法擁有非入侵性、成本低廉、即時等優點，所以常廣泛用於婦產科當中；而缺點則是容易會有雜點、變形、扭曲，這是因為超音波是藉由收到回音來成像，而聲音在傳遞過程會有慢慢衰減。如越深的器官，聲音透射量就會越小，這是由於聲音會被組織吸收或散射或反射；吸收散射造成影像sign-to-noise低；反射迴盪回音造成影像變形；相同聲阻抗的不同組織造成分界難以辨別；這是超音波先天上的特性；所以我們需要受過訓練有經驗的超音波師來判斷邊界。

而除了上述的困難點，還有一個也很棘手的問題就是根據超音波師們的判斷經驗不同及手法操作不同，所得到的造影及判斷也會有所不同；所以本論文想藉由神經網路來改善這個問題，去計算出一個不因人為而變動的標準。

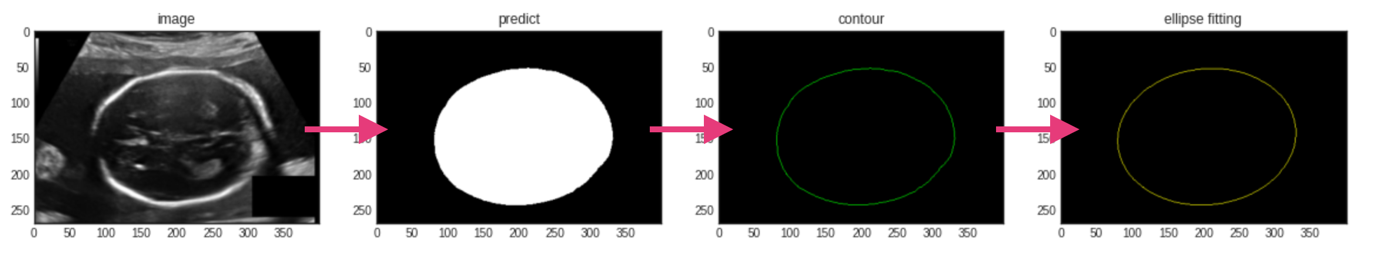
在胎兒超音波造影中有各種胎兒生物參數可以去判斷胎兒生長狀況及異常，例如：biparietal diameter (BPD)、head circumference (HC)、abdominal circumference (AC) femur length (FL) 等等。本論文所評估的胎兒參數是HC這是由於他是其中最明顯且好估計的指標。

近幾年隨著神經網路的崛起，人們開始把神經網路用臨床研究上，在之中CNN對於圖片有更良好的分割結果，如：FCNs，UNet，PSPNet；其中Unet最廣泛應用於醫療影像，它的特色是編碼在解碼的架構，並用跳躍連接讓解碼層考慮到因為降採樣而消失的細部特徵。之後Unet++在被提出，他改變了Unet跳躍連接的方式，改用了密集的跳躍連接，大大改善提升了UNet的各項分割指標，達成類似Ensemble的方式，但是由於改採用密集的跳躍連接使得需訓練的參數大幅增加，增加了訓練成本，因此Unet3+被提出。

Unet3+也改變了跳躍連接的部分，在UNet++中的跳躍連接的深層解碼層的特徵圖並沒有考慮(併接)到較淺層的編碼層所輸出的特徵圖，所以無法去融合淺層特徵考慮到細部特徵，Unet3+改善的地方就是在此，在每一個解碼層，都併接到所有編碼層的特徵圖，達到多尺度的考量，並切需訓練的參數量也比Unet++更少，在分割結果上則比Unet++稍微提升。

本論文所採用的方法是Unet3+ 加上 Attention Gate，得到精確的分割再使用最小二乘橢圓擬合得到HC估算結果，在DICE 係數上的分數為97.89 ± 1.39 %、Hausdorff Distance為1.29 ± 0.83 mm、Absolution Difference為 1.90 ± 1.80 mm、  
Difference為 -0.51 ± 2.57 mm。

方法：



流程圖如Fig.1 透過本論文提出的ATUnet3+取得胎兒頭部ROI，然後抽取邊緣再用橢圓擬合得到最終輪廓。

# Unet3+介紹

Unet3+由Huimin Huang 等人提出，將編碼層各個尺度的特徵圖，跳躍併接到每個解碼層，讓解碼層節點可以考慮多尺度的特徵，輸出較好的分割結果，  
Unet3+另一個特點是訓練參數少於unet和unet++可以更快收斂及運行速度更快。

Unet 解碼層所需訓練參數計算：

(1)

K: 卷積核大小，一般來說常用的是3

d: 的channels數

Unet++解碼層所需訓練參數計算：

(2)

:是中間層的節點。

Unet3+解碼層所需訓練參數計算：

(3)

由於Encoder的結構一樣，所以需要的訓練參數量是一樣的，主要在於Decoder的差別，所以我們計算他們的Decoder所需訓練的參數就好。

那差別主要就是channel數，透過以上公式計算每一層解碼層的節點相加，可以發現

Unet3+比Unet和Unet++所需訓練參數更低，這是由於Unet3+的解碼層節點得channel數都是固定320，而Unet和Unet++都是32 channels數

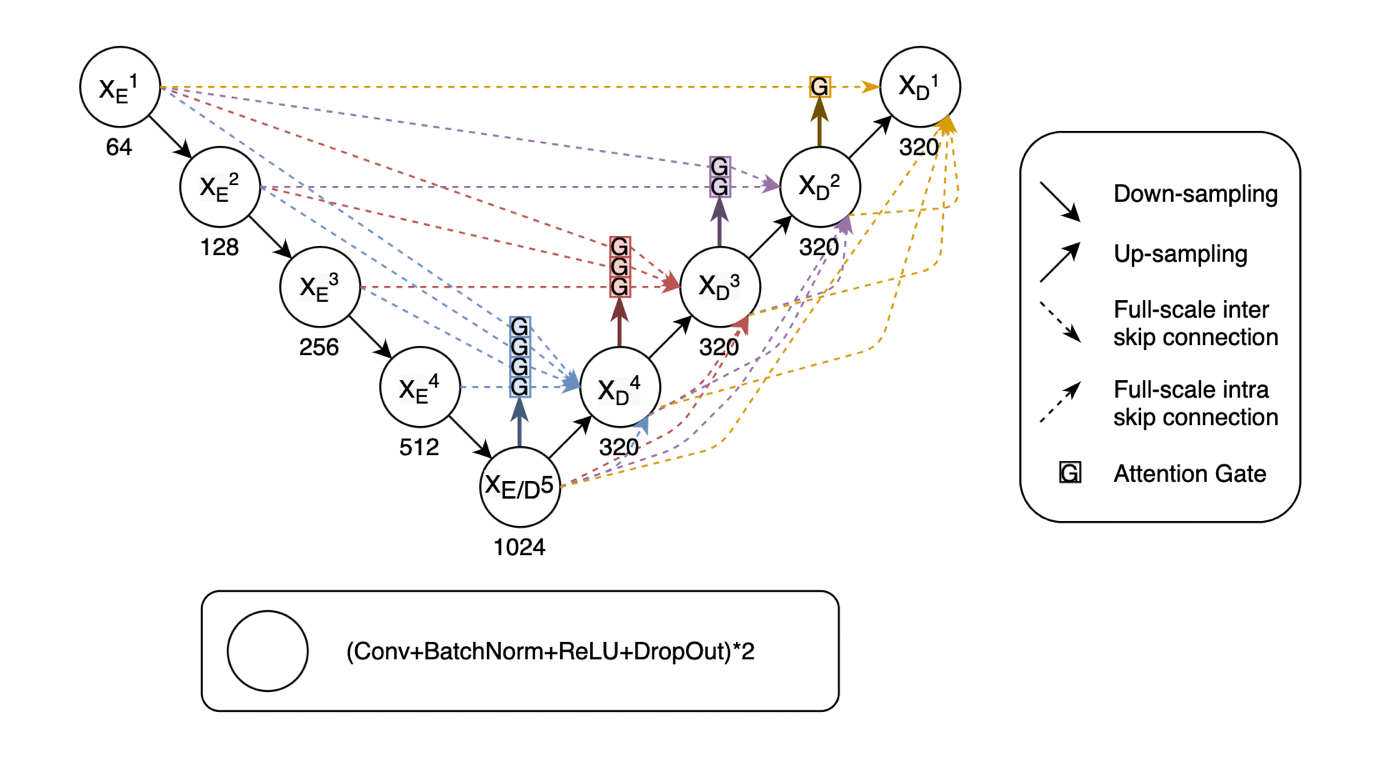
# Attention機制介紹

# 

Attetion gate 方式由JoSchlemper等人提出，該方法能讓機器給特徵圖中重要區域的比較大的權重，抑制不相關的區域，因為一般跳躍連結(Unet)是直接和解碼層做合併，沒有去學習挑選重要的部分Fig.2。

# ATUnet3+介紹

結合以上兩點，我們提出了ATUnet3+(Attetion Unet3+)的架構Fig.3，改善多尺度特徵圖的跳躍連接直接到解碼層，讓機器去學習這些特徵圖哪些部份是要著重的，從而  
達到不錯的分割結果。



ATUnet3+解碼層的節點輸出結構的公式由以下得出

(1)

通過(1)可以得到每個層解碼的輸出，

：是第i層的解碼層輸出，N：是總共有多少個解碼器

H(·)是對特徵圖做併接，G<·>是通過注意力機制後輸出，C(·)是做卷積操作，D(·)是做下採樣動作，U(·)是做上採樣動作。

# Loss function介紹

我們這裡採用混合loss funtion 來評估，結合BCE + MS-SSIM + DICE三種Loss function

分別代表，像素層級、片段層級、重疊層級，各方面來幫助我們訓練。

MS-SSIM(Multi Scale Structural Similarity index measure)

SSIM是對兩張圖的相似度做評估的指標，綜合了亮度、對比度、結構方面，但是由於圖像特徵的空間域在統計上是非平穩過程，還有人類再看一張圖時，我們都會看一小部分的區域，才能看得清晰，所以實際上會使用MSSIM。

MSSIM是用個滑動視窗NxN大小，計算區域的SSIM，對每個區域都滑過後，取平均得到，

MS-SSIM就是多尺度版本的MSSIM，會將要比較的圖像縮半，並對每個尺度都有個權重，在此根據[論文]我們設權重為[0.0448, 0.2856, 0.3001, 0.2363, 0.1333]，用5個尺度。

MS-SSIM的範圍為[0,1]，值越接近1越好，而這裡我們要算的是Loss function，所以1-(MS-SSIM)。

SSIM亮度(luminance) 評估公式：

SSIM 對比度(contrast) 評估公式：

SSIM 結構 (contrast) 評估公式：

為x的平均值，為y的平均值，為x得標準差，為y的標準差，

爲共變異數，為，為，為

設為0.01，設為0.03，為像素值的範圍，這裡設為255。

綜合以上得到SSIM公式：

可以自己設定每個函式的權重，這裡都設為1

MSSIM是取SSIM每個滑動視窗的平均，利用NxN大小的滑動視窗滑完整張圖，得到總數為M個視窗，公式為：

MS-SSIM透過縮小一半比較的圖片，可以得到不同尺度的相似程度，其公式為：

M為1，代表沒有縮小，通常對亮度相似度計算只會計算一次。

通常令

論文給出了5個尺度適合的權重

BCE(Binary Cross Entropy)

將所有像素點預測的機率和超音波師的標註的像素做評估，因為我們只有兩種分類，不是頭部就是背景，所以採用BEC，他的公式是

x爲預測圖像每個像素是不是胎兒頭部的機率，[0,1]之間

y為超音波師標注的正確機率，不是0就是1

DICE係數(Sørensen–Dice coefficient)

Dice能用來評估我們預測的圖像與超音波師的標注的重疊率。

x爲預測圖像每個像素是不是胎兒頭部的機率，[0,1]之間

y為超音波師標注的正確機率，不是0就是1

Dice係數範圍為[0,1]之間，越接近1代表我們預測的分割圖和超音波師的標註重疊率越高，因為這裡loss是要最小化，所以= (1-Dice)

# 結合上面3種loss function得到我們的混合loss function

實驗及結果