碩士論文

結合注意力機制與反向跳躍連接Unet3+神經網路用於胎兒參數測量

林芃葦 撰

110

9

國立臺北教育大學

資訊科學所

林芃葦

Peng-Wei Lin

指導教授：劉遠楨博士

Advisor : Yuan-Chen Liu, Ph. D.

中華民國 110 年 9 月

September 2021

國立臺北教育大學理學院資訊科學系

碩士論文

Department of Computer Science

College of Science

National Taipei University of Education

Master’s Thesis

結合注意力機制與反向跳躍連接Unet3+神經網路

用於胎兒參數測量

Reversed Skip Connected Unet3+ with Attention Mechanism to Calculate Fetal Biometry

摘要

本論文研究一套神經網路模型，來輔助醫生能從超音波胎兒造影裡得到胎兒指標之一的Head Circumference，幫助醫生判斷胎兒生長健康的情況。

由於超音波比其它的造影方法擁有非入侵性、便宜、即時的特性，所以被廣泛用於胎兒婦產科上，醫生可以藉由超音波來評估胎兒的生長情況。而由於一些超音波先天上的特色(缺點)，此檢查過程需要有經驗豐富的醫生操作才能判斷準確。

然而訓練一位經驗有素的超音波師需要時間和成本，根據世界衛生組織的報告，比較偏鄉或是發展中的國家普遍缺乏這些超音波師或醫師，很多母親都是在沒有超音波監控下生產，有99%生產死亡都是發生在發展中國家；所以我們希望有自動化測量來降低人力所需，並且解一同解決由於不同操作人員測量甚至同一位超音波師每次測量結果不同的這個問題。

在胎兒超音波影像裡，有各種評估胎兒健康狀況的指標，本論文選擇測量指標是Head Circumference (HC)，我們提出RevUnet3+方法，藉由修改過去文獻中的Unet3+多尺度模型，反轉跳躍連接的方式，並結合注意力機制來達到精準的胎兒頭部分割，得到頭部分割後，再用最小二乘法擬合橢圓，算出最後HC的輸出，此方法能達到媲美超音波專家的水準。

本論文測試了Unet、Unet++和Unet3+，最後是我們提出的RevUnet3+與其他研究該議題的論文，針對HC的神經網路比較。

最終，我們提出的RevUnet3+在Dice係數上的分數為97.88 ± 1.16%、Hausdorff Distance為11.25 ± 0.69 mm、Absolution Difference為 1.72 ± 1.60 mm、Difference為 0.04 ± 2.35 mm，在HC Grand Challenges 挑戰賽上獲得第一。

關鍵字：神經網路、U形神經網路、超音波造影、胎兒參數、語意分割。

Abstract

This paper is about using neural network method to calculate fetal head circumference aiding doctors and sonographers to diagnose fetal health.

Because ultrasound is safer than others medical radiography which have been used on obstetrics and gynecology extensively. However, ultrasound have some congenital feature affecting image quality, we need experienced doctors to operate accurately and get acceptable measure. But training experienced doctors need to consume a lot of cost, the lower resolution country or rural areas could not burden it. According to World Health Organization report, most death of pregnancy occur in those area that up to 99%. Therefore, we hope we can develop an automatic system for reducing the cost and unifying different measure result which got by different doctors.

In the field of ultrasonography of fetus, we get a variety of fetal biometry to assess fetus’s health. We will calculate Head Circumference (HC) as our target biometry in the paper. We propose a method to get HC segmentation which is called RevUnet3+. It modified Unet3+ by reversing skip-connection parts and adding attention-gate. After segmenting, we employ least square fitting in segmentation to get ellipse contour and calculate HC. This method is comparable with sonographer.

The paper compare some segmentation neural network model including Unet、Unet++、Unet3+ and others neural method which is proposed to calculate HC in their paper.

In conclusion, our performance score of RevUnet3+’s on Dice is 97.88 ± 1.16%, the Hausdorff Distance is 11.25 ± 0.69 mm, the Absolution Difference is 1.72 ± 1.60 mm and the Difference is 0.04 ± 2.35 mm, which is best than others method and take first place in global HC calculating contest named HC Grand Challenges.

Keywords: Neural Network、Unet、Ultrasound Image、Fetal Biometry、Semantic Segmentation.

# 目錄

[摘要 i](#_Toc81831713)

[Abstract ii](#_Toc81831714)

[目錄 iii](#_Toc81831715)

[表目錄 v](#_Toc81831716)

[圖目錄 vi](#_Toc81831717)

[1 緒論 1](#_Toc81831718)

[1.1 超音波的特色及缺點 1](#_Toc81831719)

[1.2 超音波應用於胎兒檢查 2](#_Toc81831720)

[1.3 過去機器學習的方法 3](#_Toc81831721)

[1.4 現在類神經網路的方法 3](#_Toc81831722)

[1.5 研究目的及結果 4](#_Toc81831723)

[2 文獻探討 5](#_Toc81831724)

[2.1 Unet家族 5](#_Toc81831725)

[2.1.1 Unet介紹 5](#_Toc81831726)

[2.1.2 Attention Unet介紹 8](#_Toc81831727)

[2.1.3 Unet++介紹 9](#_Toc81831728)

[2.1.4 Unet3+介紹 10](#_Toc81831729)

[2.2 注意力機制 12](#_Toc81831730)

[2.3 損失函式 13](#_Toc81831731)

[2.3.1 BCE(Binary Cross Entropy)介紹 13](#_Toc81831732)

[2.3.2 FL(Focal Loss)介紹 13](#_Toc81831733)

[2.3.3 GHM(Gradient Harmonizing Mechanism)介紹 14](#_Toc81831734)

[2.3.4 SSIM(Structural Similarity index measure)介紹 16](#_Toc81831735)

[2.3.5 MSSIM(Mean SSIM)介紹 17](#_Toc81831736)

[2.3.6 MS-SSIM(Multi Scale SSIM)介紹 17](#_Toc81831737)

[2.3.7 DSC(Sørensen–Dice coefficient)介紹 18](#_Toc81831738)

[2.4 橢圓擬合 19](#_Toc81831739)

[3 研究方法 21](#_Toc81831740)

[3.1 方法流程 21](#_Toc81831741)

[3.2 本論文提出的RevUnet3+介紹 23](#_Toc81831742)

[3.3 混合損失函式 25](#_Toc81831743)

[3.3.1 GHM loss應用於胎兒頭部分割 25](#_Toc81831744)

[3.3.2 MS-SSIM 應用於胎兒頭部分割 26](#_Toc81831745)

[3.3.3 DSC應用於胎兒頭部分割 28](#_Toc81831746)

[3.3.4 提出的混合損失函式 29](#_Toc81831747)

[3.4 計算HC值 30](#_Toc81831748)

[4 實驗與結果 31](#_Toc81831749)

[4.1 資料集來源與前處理 31](#_Toc81831750)

[4.2 資料增量 32](#_Toc81831751)

[4.3 RevUnet3+注意力機制可視化 33](#_Toc81831752)

[4.4 評估指標介紹 35](#_Toc81831753)

[4.4.1 DSC(Dice Coefficient)指標 35](#_Toc81831754)

[4.4.2 HD(Hausdorff distance)指標 35](#_Toc81831755)

[4.4.3 DF(Difference)指標 36](#_Toc81831756)

[4.4.4 ADF(Absolute Difference)指標 36](#_Toc81831757)

[4.5 實驗結果比較 37](#_Toc81831758)

[4.5.1 RevUnet3+ 三個期孕期指標分數 37](#_Toc81831759)

[4.5.2 HC Grand Challenge 排名及平均比較結果 37](#_Toc81831760)

[4.5.3 經典分割網路與其他研究該HC議題論文提出的網路比較結果 37](#_Toc81831761)

[4.5.4 與Unet家族視覺化比較分割最差的造影 39](#_Toc81831762)

[4.6 訓練超參數設定與執行環境 40](#_Toc81831763)

[5 結論與未來發展 41](#_Toc81831764)

[參考資料 42](#_Toc81831765)

# 表目錄

[表1 RevUnet3+三個孕期的指標分數 37](#_Toc81832719)

[表2 RevUnet3+與挑戰賽比較成果 37](#_Toc81832720)

[表3其他論文及經典分割網路和我們的比較結果 38](#_Toc81832721)

[表4超參數設定與執行環境 40](#_Toc81832722)

# 圖目錄

[圖1 超音波先天上的缺點 1](#_Toc81832487)

[圖2 Unet架構圖 5](#_Toc81832488)

[圖3 胎兒頭部輪廓從淺層至深層碼層所輸出的特徵圖 7](#_Toc81832489)

[圖4 AttentionUnet架構圖 8](#_Toc81832490)

[圖5 Unet++架構圖 9](#_Toc81832491)

[圖6 Unet3+架構圖 10](#_Toc81832492)

[圖7 注意力機制流程 12](#_Toc81832493)

[圖8 研究方法偽代碼 21](#_Toc81832494)

[圖9 研究方法流程圖 22](#_Toc81832495)

[圖10 每個流程的視覺化 22](#_Toc81832496)

[圖11 RevUnet3+架構圖 23](#_Toc81832497)

[圖12 GHM Loss 偽代碼 25](#_Toc81832498)

[圖13 Ms-ssim Loss 偽代碼 27](#_Toc81832499)

[圖14 DSC Loss偽代碼 28](#_Toc81832500)

[圖15訓練模型偽代碼 29](#_Toc81832501)

[圖16調整原始資料標註 31](#_Toc81832502)

[圖17資料增量例子 32](#_Toc81832503)

[圖18測試注意力機制的原始資料 33](#_Toc81832504)

[圖19各層注意力機制所關注的地方，由上往下10個epochs的變化 34](#_Toc81832505)

[圖20我們提出的網路與Unet家族比較結果 39](#_Toc81832506)

# 緒論

## 超音波的特色及缺點

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、偽影，這是因為超音波是藉由接收回音來成像，而聲音在傳遞過程中會慢慢衰減或是被身體組織反射、散射、折射，所以如果機器收到反射的量太低、消失或是收到其他組織的散射時，會導致成像出現上述缺點。

如吸收散射造成影像sign-to-noise低；反射迴盪造成影像變形偽影；相同聲阻抗的不同組織成像的分界難以辨別；這是超音波先天上的特性，圖1表示在胎兒頭部影像中常出現的問題，讓判斷便困難，(a)有時會照到母親的胎盤或是臍帶混淆判斷，(b)(d)因聲音傳播散射、折射等等回傳至超音波接收器導致模糊對比度低，(c)超音波的反射有些沒有反射至接收器。

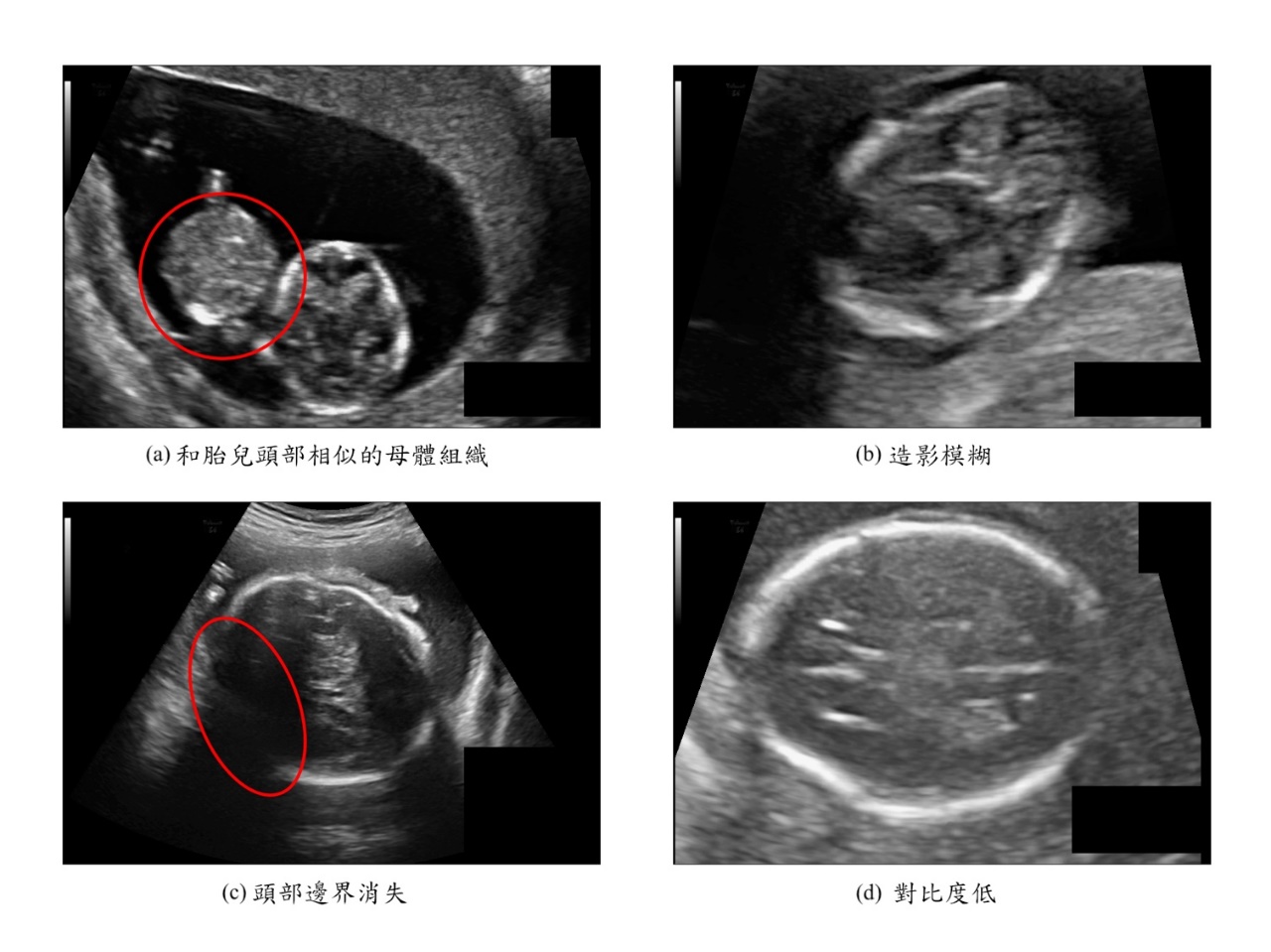


圖1超音波先天上的缺點

而除了上述的困難點，還有一個也很棘手的問題就是測量結果會依據超音波師們的判斷經驗不同及手法操作不同，所得到的造影和評估結果也會有很大的不同，所以我們也需要一套標準統一的測量方法。

## 超音波應用於胎兒檢查

超音波廣泛用於婦產科上，我們可以藉由超音波來評估胎兒的生長情況，但由於超音波先天上的缺點1.1節所述，需要有經驗的醫生或超音波師判斷，可是訓練一位經驗有素的超音波師需要時間和成本，比較偏鄉或是發展中的國家普遍缺乏，很多母親都是在沒有超音波監控下生產，而有99%生產死亡都是發生在發展中國家[1]；加上每位超音波師的測量結果都會有所不同，甚至同一位超音波師在同一天不同時段的測量也會有所差異，所以能夠自動化取代人力及統一標準的測量超音波胎兒的研究是非常重要的議題，過去到現在都有很多文獻在研究。

在胎兒超音波造影中有各種胎兒生物參數可以去評估胎兒生長狀況及異常，例如：Biparietal Diameter (BPD)、Head Circumference (HC)、Abdominal Circumference (AC)、Crown-Rump Length (CRL) 和Femur length (FL) 等等。

HC、BPD和CRL是用來評估胎兒大小和胎齡的指標，可用來判斷胎兒生長情況有無異常，他們不同之處是測量不同孕期時的精準度，CRL在胎兒6週到13週時測量是最精準的指標，而之後CRL誤差會增大[2]；所以13週之後會改用BPD和HC指標，BPD是14週到36週最精準的指標；而最後則是HC，36到40週最精準，HC可用來預測胎兒預產期。

本論文所選擇評估的胎兒參數是HC，也就是測量胎兒頭部周長，當然同時也可測量到BPD，也就是HC的短軸長度，但是因為資料集的標注只有HC周長，所以我們選擇測量HC。

## 過去機器學習的方法

過去已有使用傳統的機器學習方法來自動化偵測胎兒HC，但都有些缺點，例如G. Carneiro等人[3]使用Probabilistic boosting tree去預估和分類超音波分割結果，雖然結果不錯，不過缺點是要訓練樹需要有大量專家標記完好且複雜的資料；J. Li等人[4]使用隨機森林去定位胎兒頭部位置然後使用Phase symmetry 和去擬合橢圓計算HC，但是這方法需要先知道胎齡及超音波掃描的深度；W. Lu等人[5]採用 K-mean 和 Morphological operation 來取得頭部片段，然後用 Iterative randomized Hough transform 去擬合頭部橢圓，但問題是K-mean不能有效的抽出頭部片段，這是由於超音波的雜訊太多，讓後續的 Hough transform 不穩定；R. V. Stebbing等人[6]提出使用 boundary fragment model 方法再丟給隨機森林做邊緣偵測，然而這個方法會因為超音波造影有其他解剖構造，例如胎盤的影像強度和胎兒頭部相似而影響此方法的穩定性。

## 現在類神經網路的方法

近幾年隨著神經網路的崛起，人們開始把神經網路運用於臨床研究上，在之中由於CNN對於圖片有良好的分割結果，所以也被用於醫學影像上，如Matthew Sinclair等人[7]提出使用全卷積神經網路 FCNs來測量胎兒HC，Zahra Sobhaninia等人[8]提出使用多尺度輸入的Link-Net來測量胎兒HC，Xing Yanyan等人[9]則使用融合型Unet++，在結果顯示上，在各指標都超越了舊有機器學習的方法，達到不錯的分數。

目前這些用於醫學影像的神經網路架構都是採用類似Unet[10]的架構的思想，本論文也使使用此架構做修改。

這邊簡略介紹一下Unet的特色，Unet使用類似Auto-encoder[11]的編碼在解碼方式，並使用跳躍連接來讓解碼層考慮到因為降採樣而消失的細部特徵，Unet在2015年在ISBI的黑腹果蠅第一齡幼蟲腹神經索的分割挑戰賽中排名第一，此後大量有關醫學影像分割的問題都會採用Unet的方式。

在Unet之後，Zongwei Zhou等人[12]改良了跳躍連結的部分，提出Unet++，採用密集的跳躍連接，達成類似集成子Unet的方式，大大提升了Unet的各項目分割指標；而代價是由於密集的跳躍連接使得需訓練的參數大幅增加，增加了訓練成本和收斂速度，Xing Yanyan 等人[9]提出的融合型Unet++測量HC，就是基於Unet++上多增加了一層捲積把子Unet的各輸出結果融合才做最後的輸出。

而除了上述訓練成本的問題，Unet和Unet++的跳躍連結的每一個解碼層只有使用同一層的編碼層，而沒有考慮到其他尺度的編碼層，無法獲得更豐富多尺度的編碼層資訊，因此Huimin Huang等人[13]提出了Unet3+。

Unet3+改變了跳躍連接的部分，讓每一個編碼層的特徵圖，都會併接到解碼層，除此之外，解碼層也會考量過去解碼層的特徵圖，達到多尺度的考量，並且由於固定解碼層的通道數，需訓練的參數量也比Unet++和Unet少，達到收斂快速及防止過擬合的效果，在分割結果上則比Unet++稍稍提升；但是該方法的多尺度跳躍連接的降採樣，與編碼層的降採樣部分都採用同樣的Maxpool方式，損失的邊緣細節也會相同。

## 研究目的及結果

本論文所提出的方法是對於Unet3+多尺度跳躍連結部分保持疑問，修改跳躍連接的方式後再引入注意力機制成為RevUnet3+，採用混合Loss Function做訓練，結合GHM loss、Ms-ssim loss 和 Dice loss三種，分別代表像素層級、片段層級和分割層級；得到分割結果後，再使用最小二乘橢圓擬合，去計算出HC，在HC Challenge比賽上獲得第一的佳績。

最後，使用Dice係數、Hausdorff Distance、Absolution Difference、Difference作為評估的指標，對Unet、Unet++、Unet3+與其他研究該議題的論文跟我們提出的RevUnet3+做比較。

# 文獻探討

## Unet家族

Unet於2015年提出，在近年的醫學影像被大量使用，後續的人們幾乎都使用此架構上改進，如ResUnet、AttentionUnet、LinkNet、R2Unet等等，本論文接下來所介紹的是我們靈感來源的Unet。

### Unet介紹



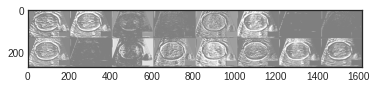
圖2 Unet架構圖

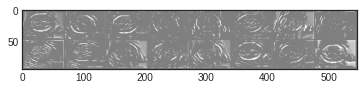
Unet由Olaf Ronneberger[10]等人提出，是針對只有少量資料的醫學影像的而發明的模型，特色是收縮(編碼)、擴展(解碼)及跳躍連接；Unet在2015年在ISBI的黑腹果蠅第一齡幼蟲腹神經索的分割挑戰賽中排名第一，圖2為Unet的架構。

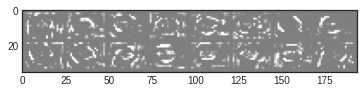
Unet概念受到Auto-encoder[11]影響，只不過他們的目的不同，Unet是要輸出分割，Auto-encoder是學習如何壓縮圖片，讓壓縮完的圖片在復原後也能和原圖片相似且少失真，兩者的架構很相似，如Unet的收縮(contracting)和擴展(expansive) 對應Auto-encoder的編碼(encoder)和解碼(decoder)。

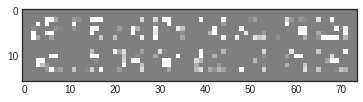
Unet還多了一個特色，就是加上了跳躍連接的部分，能把收縮中的節點輸出(特徵圖)去併接到擴展節點的特徵圖，如此一來擴展節點能考慮到收縮部分因為降採樣而消失的圖片細節。

Unet中，越深層的節點能收到越抽象高級的資訊如位置、片段或方向，而淺層的節點收到的資訊通常是低層次的，如邊緣，線條，等等簡單的資訊，圖3為同一張超音波造影的編碼層特徵圖輸出，每層各只取16張特徵圖，由淺到深。









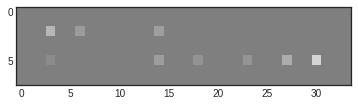


圖3胎兒頭部輪廓從淺層至深層碼層所輸出的特徵圖

每個節點為兩次捲積再輸出，加上BatchNorm能對每個特徵圖偏移的分佈做修正成正態分佈，解決了偏移所造成的神經網路收斂變慢[14]。

### Attention Unet介紹

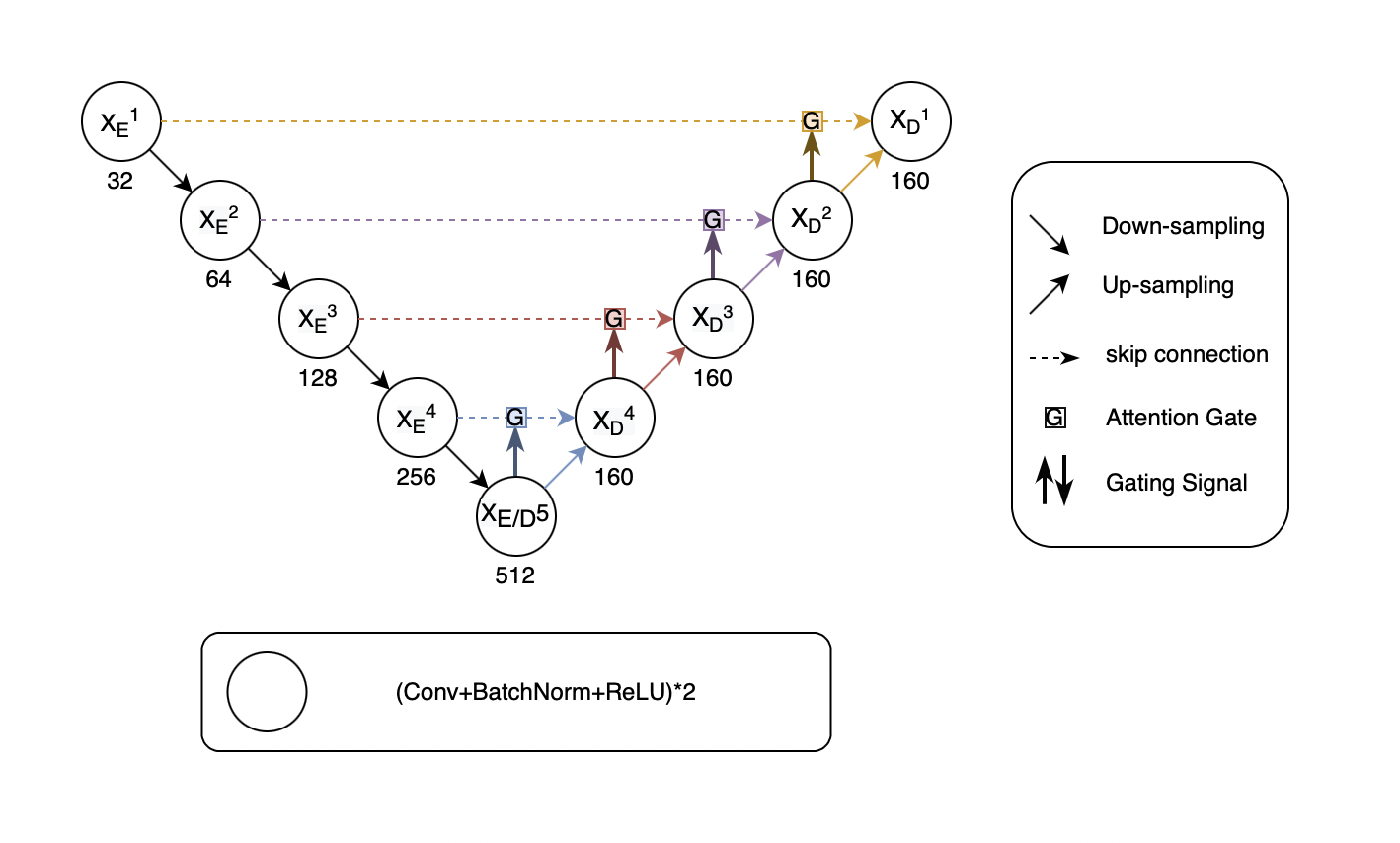


圖4 AttentionUnet架構圖

Attention Unet於2018年由Ozan Oktay[15]提出，架構他改善了Unet經由跳躍連接直接併接的特徵圖無法關注重要部分的問題，使用Attention Gate對這些特徵圖做權重計算，這些權重是經由機器自己學習的，而之所以用深層的節點來監督淺層的特徵圖的原因是因為深層節點抽取的資訊較重要，我們只要取得根據深層節點附近的淺層節點資訊就好；圖4為Attention Unet架構圖。

### Unet++介紹

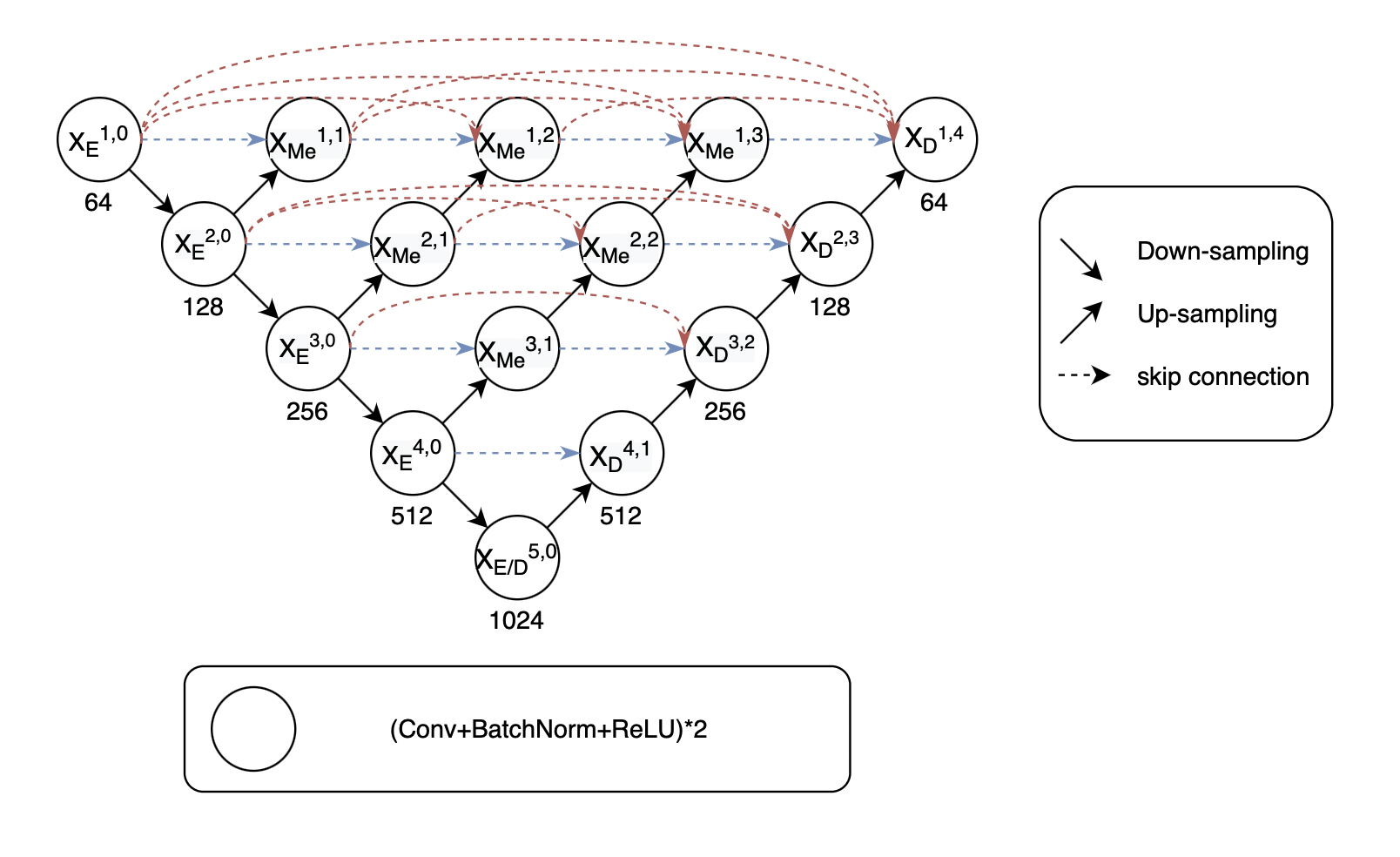


圖5 Unet++架構圖

Unet++由Zongwei Zhou[12]等人提出，可以看做Unet的父集，提出理念是因為抱持著對於Unet到底要用幾層才能擁有最好的精確度的疑問，那麼乾脆把多個不同層數的Unet互相連結合而為一，誕生出Unet++，圖5為Unet++架構。

Unet++主要改良了跳躍連接的部分，採用了密集的跳躍連接，所以擁有較多層數的Unet可以藉由連結幫助比較淺的Unet的權重訓練，而較多層數的Unet也能利用連結得到較淺Unet的特徵圖。

最後作者還提出剪枝的想法，加上了深監督的後，能讓每個不同層數的Unet都能輸出分割圖，這樣就可以在未來根據訓練參數的成本和精準度提升程度，來權衡要使用多少層的Unet就好，剪去提升不大的Unet節點。

### Unet3+介紹

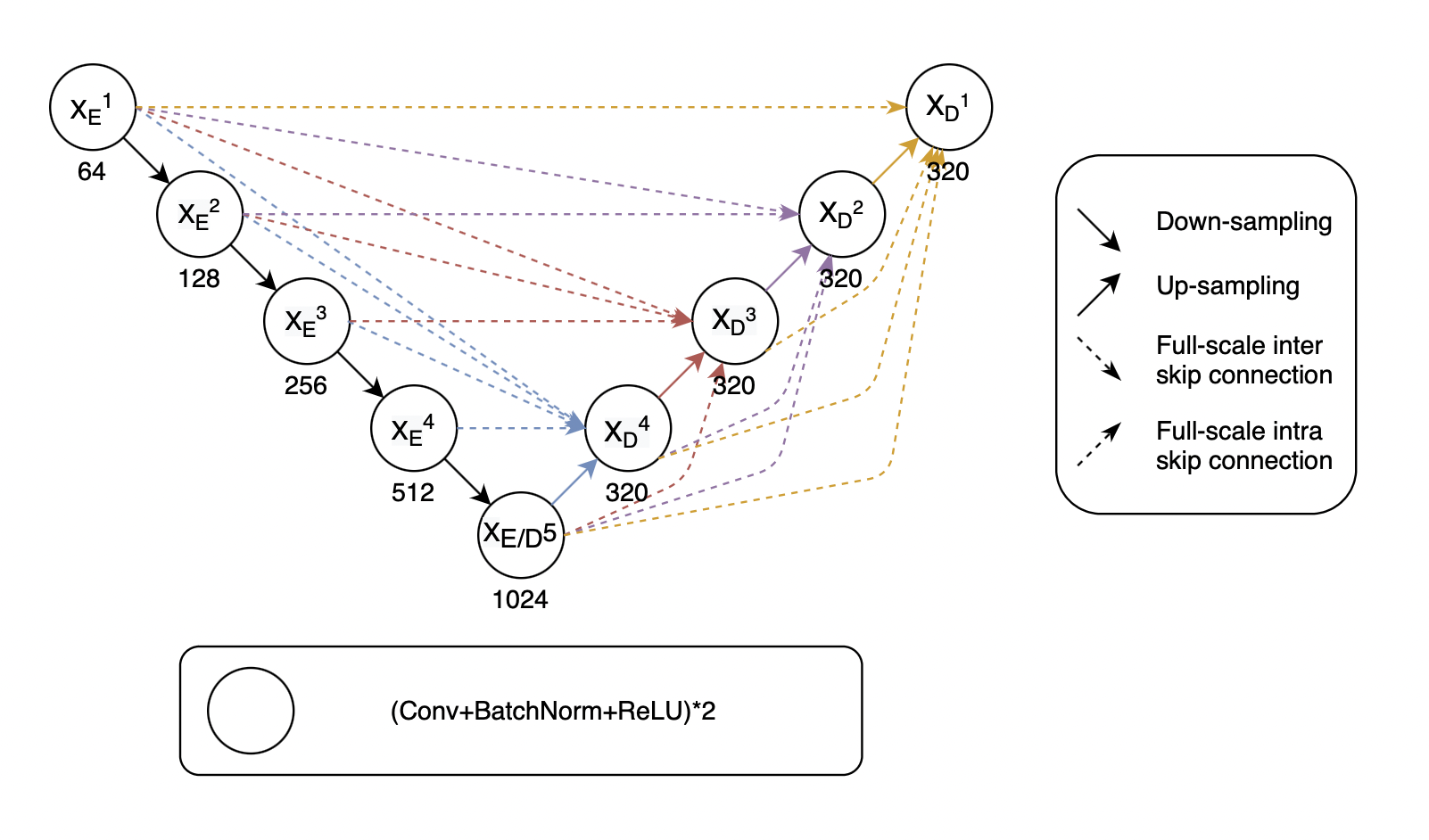


圖6 Unet3+架構圖

Unet3+由Huimin Huang[13]等人提出，特色是將編碼層各個尺度和深層的解碼層的特徵圖，跳躍併接到每個解碼層，讓解碼層節點可以考量到多尺度的特徵，來輸出較好的分割結果，而且訓練參數少於Unet和Unet++，可以達到更快收斂及運行速度更快的效果，圖6為該架構。

對於Unet3+比Unet和Unet++訓練參數還少的原因，可由以下解碼層所需要公式算出。

Unet 解碼層所需訓練參數計算：

|  |  |
| --- | --- |
|  | (1) |

K：卷積核大小，本論文用的數值是3。

d： 的channels數。

Unet++ 解碼層所需訓練參數計算：

|  |  |
| --- | --- |
|  | (2) |

：是中間層節點。

Unet3+ 解碼層所需訓練參數計算：

|  |  |
| --- | --- |
|  | (3) |

只考慮Decoder的原因是因為他們Encoder的結構一樣，所以Encoder的訓練參數量是一樣的。

從式子上來看，可以看到Unet++(2)因為比Unet**Error! Reference source not found.**多了去併接解碼層的部分，所以Unet++的訓練參數大於Unet，而Unet和Unet3+參數差距主要原因就是channel數，透過(3)可以發現Unet3+的解碼層都是固定320，而Unet是 channels數指數成長。

## 注意力機制



圖7 注意力機制流程

注意力機制常用於自然語言處理、知識圖譜、影像分析，如image caption[16]、機器翻譯[17]、分類任務[18][19]等等，而在醫學影像分割上，Ozan Oktay等人在Unet上應用了注意力機制[15]，該注意力機制用於跳躍連接中，讓機器要併接的特徵圖中給重要區域比較大的權重，抑制不相關的區域，由於深層的節點能獲得更多高級抽象特徵訊息，所以用深層的節點來監督上一層編碼層節點的輸出，圖7為注意力機制的流程，注意力公式為(4)，輸出為加權後的特徵圖：

|  |  |
| --- | --- |
|  | (4) |

：編碼層輸出。

：解碼層輸出。

：1×1卷積，用來讓和的channels數目一樣，才可以做矩陣相加。

：Relu激活函式。

：Sigmoid激活函式。

：Hadamard乘積。

## 損失函式

在神經網路中，我們常採用幾種指標來評估分割的好壞，計算出和標註的誤差，

這個目的是讓機器在後傳導的過程中，學習怎麼更新網路權重讓這些誤差越小越好，我們會對這篇論文所採用的loss function簡單介紹一下。

### BCE(Binary Cross Entropy)介紹

BCE是Cross Entropy的一個特例，用來評估兩個機率分佈的差異，值越大代表混亂程度越大也就是機率分佈差距很大，值越小代表機率分佈越相近，我們希望值能越小越好，常用來當分類或是分割任務的loss function[20]。

它的公式是：

|  |  |
| --- | --- |
|  | (5) |

：為預測值的機率，。

：為實際值的機率，。

### FL(Focal Loss)介紹

雖然BCE是個好用的指標，但是對所有樣本都一視同仁，當模型收斂時，易分類的佔了一大部分，導致機器只學習簡單易分類的樣本就好，因為就算困難樣本分類很差，我只要顧好簡單樣本得分就很高。

舉個例子，有小明讀數學時，有加減題和乘除題，其中有本習題本有90%都是加減題，只有10%是乘除題，老師只告訴小明把這習題本全部做完反覆練習，那麼小明就乾脆專注於加減題，而不專注於乘除題，因為就算乘除題全錯，小明還有90分，可實際上把小明真的到了考試時，加減題和乘除題數量相當，那麼將會很慘，所以我們希望小明應該不要那麼專注於習題本的加減題，而是花時間訓練乘除題。

我們對模型訓練也是一樣道理，希望能更專注學習較困難的樣本，所以Focal loss修改了BCE一些地方，Focal loss採用的方法是，把較簡單的樣本給很小的權重，梯度乘上權重後，神經網路就比較不會專注學習那些簡單樣本，而困難樣本則反之，給大權重，讓機器學習。

這裡說明下困難簡單樣本的意思，就是看置信程度，舉個二分類例子就是，要機器分辨一張圖片，圖片不是貓就是狗，我們把一張狗圖片丟進機器判斷，機器判斷有51%是狗或是判斷錯誤成貓，那麼這張圖片是困難的，若機器判斷99%是狗，那這張圖片對機器分辨是簡單的。

將上述概念應用於BCE就是Tsung-Yi Lin[21]等人提出Focal loss，Focal loss公式為：

|  |  |
| --- | --- |
|  | (6) |

：為預測值的機率，。

：為實際值的機率，。

：根據[21]，值設為0.25。

：根據[21]，值為2。

### GHM(Gradient Harmonizing Mechanism)介紹

雖然Focal loss會讓機器更加學習困難樣本，但是困難樣本中的極困難樣本卻可能是離群值、雜訊，所以應該除了不要注意簡單樣本外，也不要那麼過分注意極困難樣本，所以Buyu Li等人[22]提出了GHM(Gradient Harmonizing Mechanism)，藉由給予極困難和極簡單樣本低的權重，讓這兩種梯度變小。

方法是先訂個判斷樣本難易度的g，越接近1則是困難樣本，0則反之。

公式為：

|  |  |
| --- | --- |
|  | (7) |

：為預測值的機率，。

：為實際值的機率，。

然後訂個區間如[0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]，藉由算出g落在哪個範圍，給越多落在該範圍的圖片很小的梯度權重，這是由於根據論文，發現模型收斂後，極困難樣本和極簡單樣本的數量通常比普通難度樣本多很多，所以根據數量來對應權重，越多數量的區間權重越低。

依照上述想法，設計了梯度密度，代表了有多少樣本數量在以長度的區間範圍，結合(9)(10)，梯度密度的公式為：

|  |  |
| --- | --- |
|  | (8) |

：為樣本數。

：第k個樣本的，。

：自己訂的區間中心，如

：為區間中心，左右的總長度。

：判斷有沒有落在區間範圍內，公式為：

|  |  |
| --- | --- |
|  | (9) |

：以為中心，算實際區間的長度，如果直接把乘2倍，可能會超出範圍，公式為：

|  |  |
| --- | --- |
|  | (10) |

代表了有多少樣本數量在以長度的區間範圍，而

因為我們要越大權重越小，所以定義的調和參數，給BCE的當權重：

|  |  |
| --- | --- |
|  | (11) |

最後結合**Error! Reference source not found.**的BCE算出GHM，公式為：

|  |  |
| --- | --- |
|  | (12) |

### SSIM(Structural Similarity index measure)介紹

SSIM[23]是對兩張圖的相似度做評估的指標，結合了亮度、對比度、結構方面來評估相似度。

SSIM亮度(luminance) 評估公式：

|  |  |
| --- | --- |
|  | (13) |

SSIM對比度(contrast) 評估公式：

|  |  |
| --- | --- |
|  | (14) |

SSIM結構 (structure) 評估公式：

|  |  |
| --- | --- |
|  | (15) |

：為機器預測圖片。

：為我們希望的目標圖片。

：x的平均值。

：y的平均值。

：x的標準差。

：y的標準差。

：共變異數。

：值為0.01

：值為0.03

：值為255，是像素8bit值的最大範圍。

：值為。

：值為。

：值為。

綜合(13)(14)(15)得到SSIM公式：

|  |  |
| --- | --- |
|  | (16) |

：可以自己設定每個函式的權重。

### MSSIM(Mean SSIM)介紹

但是由於圖像特徵的空間域在統計上是非平穩過程，以及人類再看一張圖時，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使用滑動視窗看局部區域，MSSIM是取SSIM每個滑動視窗的平均，利用N × N大小的滑動視窗滑完整張圖，得到總數為M個視窗，公式為：

|  |  |
| --- | --- |
|  | (17) |

### MS-SSIM(Multi Scale SSIM)介紹

MS-SSIM[24]就是多尺度版本的MSSIM，會將要比較的圖像縮半，分別套用MSSIM，並給他們各自的權重，這是由於人類在看一個東西時，和東西的距離也會影響我們感官，距離遠可能只看到外形，距離近則能看到細節，實現此概念最簡單的方法就是縮放圖片的尺度。

MS-SSIM的範圍為[0,1]，值越接近1代表越相似，而這裡要算的是Loss function，所以我們需要用1減去(17)：

|  |  |
| --- | --- |
|  | (18) |

MS-SSIM透過不斷縮小要比較的圖片，可以得到不同尺度的相似程度，其公式為：

|  |  |
| --- | --- |
|  | (19) |

值為設定為共多少個尺度，對亮度相似度只會計算最小尺度的亮度。

根據[24]，通常令。

### DSC(Sørensen–Dice coefficient)介紹

DSC[25]能用來評估兩張圖片的分割物體的集合相似程度，公式是：

|  |  |
| --- | --- |
|  | (20) |

：機器預測的圖像裡某分割所含的像素的機率，。

：目標分割的每個像素的機率，。

DSC的範圍為[0,1]之間，越接近1某個分割的重疊及相似程度高，而因為這裡loss是要最小化，所以我們需要用1減去(20)：

|  |  |
| --- | --- |
|  | (21) |

## 橢圓擬合

使用OpenCV的findContours()方法[26]取得的輪廓，再用fitEllipse()方法[27]得到描述橢圓的6個橢圓一般式參數，最後轉為描述橢圓標準式的5個參數。

fitEllipse()使用的演算法由Andrew W. Fitzgibbon等人[27]提出，以下為計算方法，首先橢圓一般式為：

|  |  |
| --- | --- |
|  | (22) |

我們想要讓所有輪廓點帶進去越約近零越好，方法是解最小二乘法，先把(22)轉為矩陣形式相乘：

|  |  |
| --- | --- |
|  | (23) |

|  |  |
| --- | --- |
|  | (24) |

|  |  |
| --- | --- |
|  | (25) |

求最佳解，但由於當時，有個任意的縮放因子，使得也會滿足上述條件，代表同樣的橢圓，所以把(25)限制改為，求最佳解：

|  |  |
| --- | --- |
|  | (26) |

然後求拉格朗日乘子，求得的，最後將轉換成描述橢圓的5個參數，分別依序為：橢圓中心x座標位置，橢圓中心y座標位置，半長軸，半短軸，角度，轉換公式為：

|  |  |
| --- | --- |
|  | (27) |

|  |  |
| --- | --- |
|  | (28) |

|  |  |
| --- | --- |
|  | (29) |

|  |  |
| --- | --- |
|  | (30) |

|  |  |
| --- | --- |
|  | (31) |

最後從(29)(30)得到的長短半軸來計算橢圓週長。

|  |  |
| --- | --- |
|  | (32) |

# 研究方法

## 方法流程

輸入一張超音波造影，用本論文提出RevUnet3+神經網路取得分割，使用OpenCv的FindContours函數取得最多點的輪廓作為胎兒頭部位置，最後使用OpenCv的fitEllipse擬合出橢圓，取得橢圓的半長軸和半短軸後，乘上該造影的像素大小，取得現實單位的周長(mm)，此周長為HC，圖8是演算法偽代碼，圖9流程圖，圖10每個階段的視覺化處理結果。

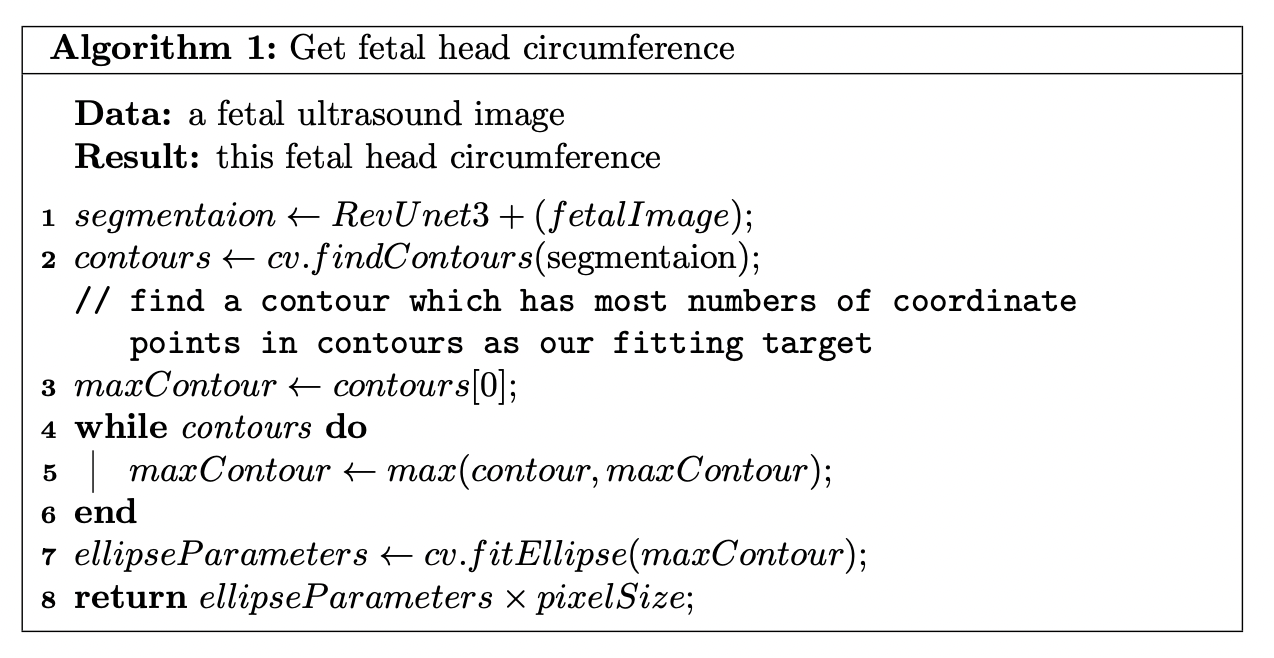


圖8 研究方法偽代碼

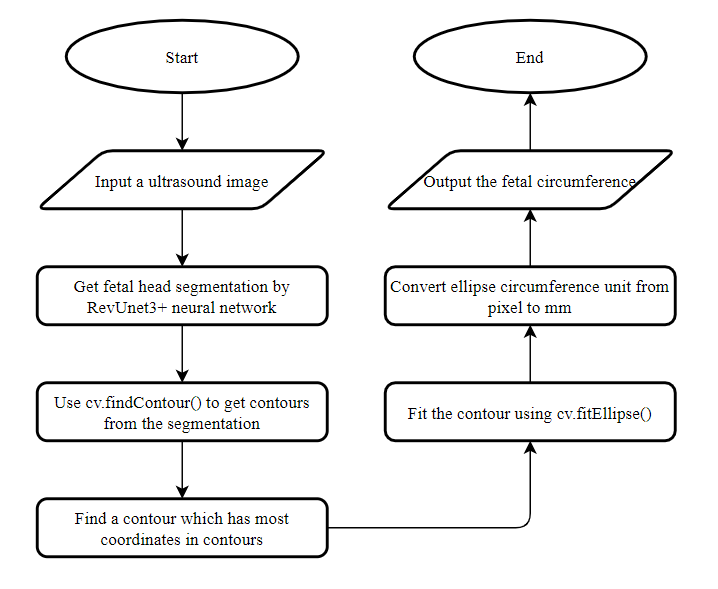


圖9 研究方法流程圖

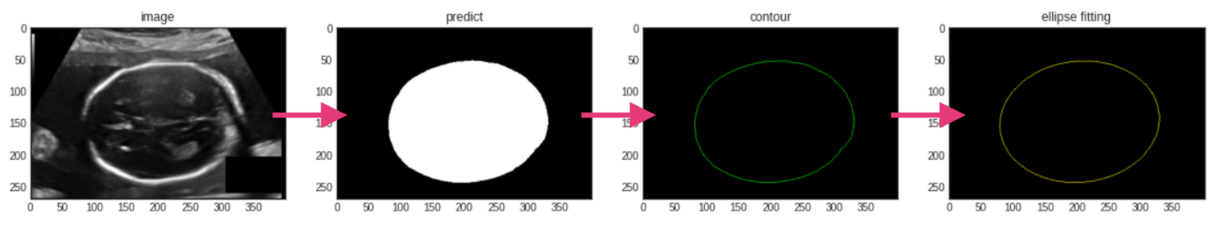


圖10每個流程的視覺化

## 本論文提出的RevUnet3+介紹

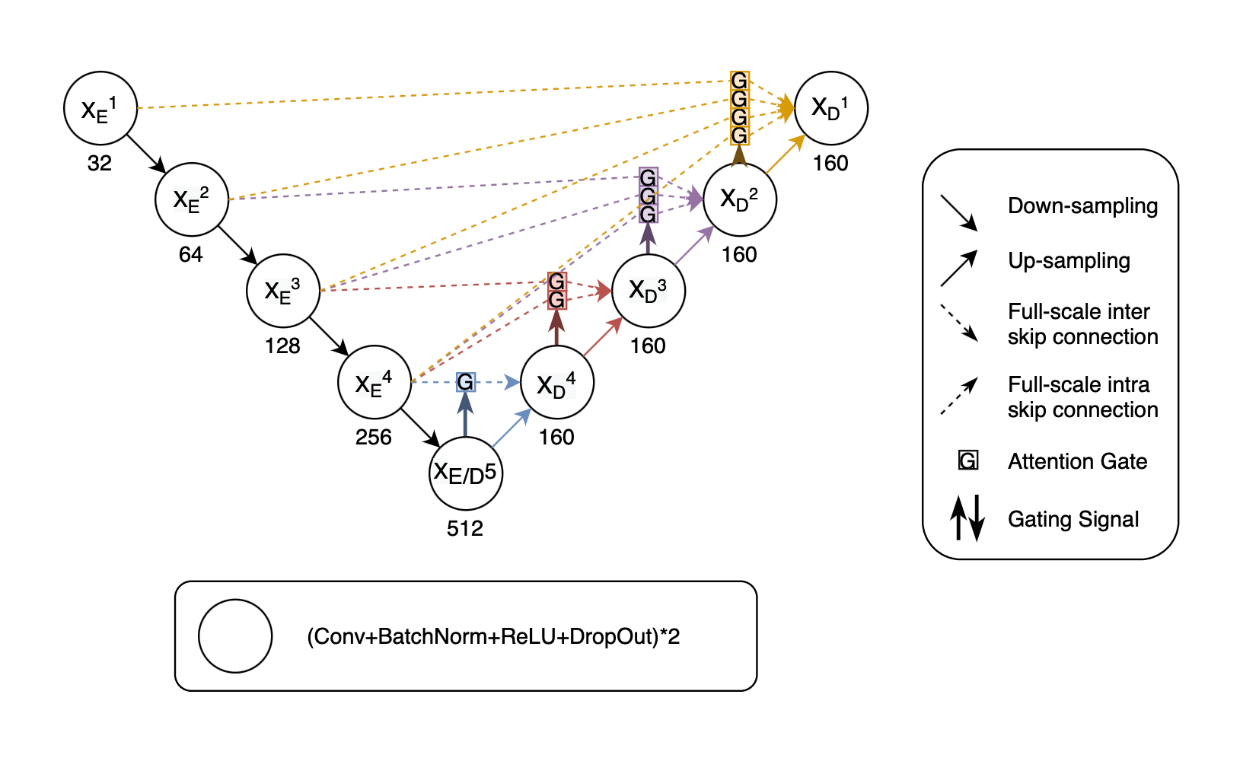


圖11 RevUnet3+架構圖

圖11為更改Unet3+跳躍連接而提出的RevUnet3+架構，理由為以下，

從圖6的Unet3+對比可以發現，我們更改了多尺度跳躍連結的部分，以倒過來的方式連結，這是因為我們觀察到Unet3+的編碼節點的特徵圖再經過多尺度跳躍連接要併接到解碼節點時，必須做Maxpool的操作才能讓特徵圖的長寬尺一致，這樣才可併接；但是這就與一般的下採樣方式一樣(黑色箭頭)，也是透過Maxpool方式，這樣兩者都會損失的一樣的細節，對此多尺度能幫助的效果上，我們保持疑問。

我們提出RevUnet3+，藉由把跳躍連接以倒過來的方式連結，這樣好處是，不會把編碼節點的特徵圖做下採樣，這樣就不會丟失細節，又能讓解碼層的節點能夠考量到多尺度的編碼節點的特徵圖。

補充一點，通常越淺層的編碼層節點的特徵圖，取得的資訊都是簡單的，如線條，邊緣，而越深層則是比較抽象複雜的資訊，如位置，片段。

最後結合注意力機制，使用在每個要併接的特徵圖上，讓機器學習判斷要併接的特徵圖哪些部分是重要的，在4.3節會有RevUnet3+每一層的注意力機制的視覺化。

另外我們對特徵圖有加上Spatial Dropout[28]，訓練時，每次迭代都會隨機把特徵圖某些channels的神經元值設為零，依靠剩下的神經元來輸出預測結果，測試時則會讓所有神經元恢復，這樣的做法能防止過擬合。

我們提出的RevUnet3+解碼層的節點輸出的公式由以下得出：

|  |  |
| --- | --- |
|  | (33) |

：是第i層的解碼層節點輸出。

：是第i層的編碼層節點輸出。

：是總共有多少個解碼層。

對特徵圖做卷積。

是對特徵圖做併接。

是通過注意力機制後輸出。

是做上採樣動作，這裡使用雙線性插值來上採樣。

由於是對少量資料訓練，為防止過擬合，我們把channels數比原先Unet家族降低一半，這樣收斂及訓練速度更快。

## 混合損失函式

我們這裡採用混合Loss function 來評估，結合了 三種Loss function，這些Loss function依序代表，像素層級、片段層級、分割層級，透過各方面來幫助我們神經網路訓練。

### GHM loss應用於胎兒頭部分割

我們用來將預測的圖片的像素點機率和超音波師的標註的像素點擊率做機率分佈的像似程度，達到像素層級的評估，而因為我們只有兩種分類，不是頭部就是背景，所以採用二元分類BCE的GHM loss：

|  |  |
| --- | --- |
|  | (34) |

：為預測圖像的所有像素點。

：我們預測的圖像的每個像素的胎兒頭部的機率，。

：為超音波師標注的每個像素的胎兒頭部的機率，。

圖12為我們在此使用的的偽代碼，我們設置10個區間難度

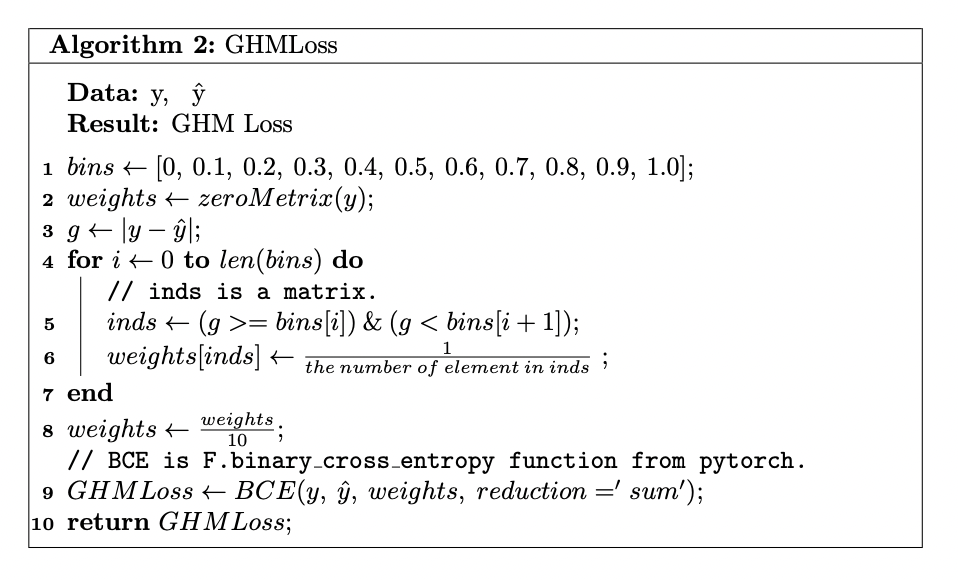


圖12 GHM Loss 偽代碼

### MS-SSIM 應用於胎兒頭部分割

藉由MS-SSIM使用滑動視窗和縮放樣本，我們能評估一張造影各尺度的各個片段的相似度，從而達到各尺度的片段層級的評估效果，我們在本論文用了5個尺度，每給尺度降採樣，縮小2倍，公式為：

|  |  |
| --- | --- |
|  | (35) |

：我們預測的圖像的每個像素的胎兒頭部的機率，。

：為超音波師標注的每個像素的胎兒頭部的機率，。

：值為5，代表五種尺度。

：亮度權重。

：對比權重。

：結構權重。

我們令，並根據論文[24]給出了5個尺度適合的亮度權重，依序是

，最後因為是要最小化，所以1減去(35)算出：

|  |  |
| --- | --- |
|  | (36) |

圖13為我們在此使用的的偽代碼，超參數設置為以下：

winSize = 11 // sliding window 大小。

winSigma = 1.5 //高斯分佈的標準差。

//穩定計算luminance、contrast及structure時分母接近0的情況。

//同上述。

L = 255 //像素的動態範圍，這裡是用8-bit灰階，所以是255。

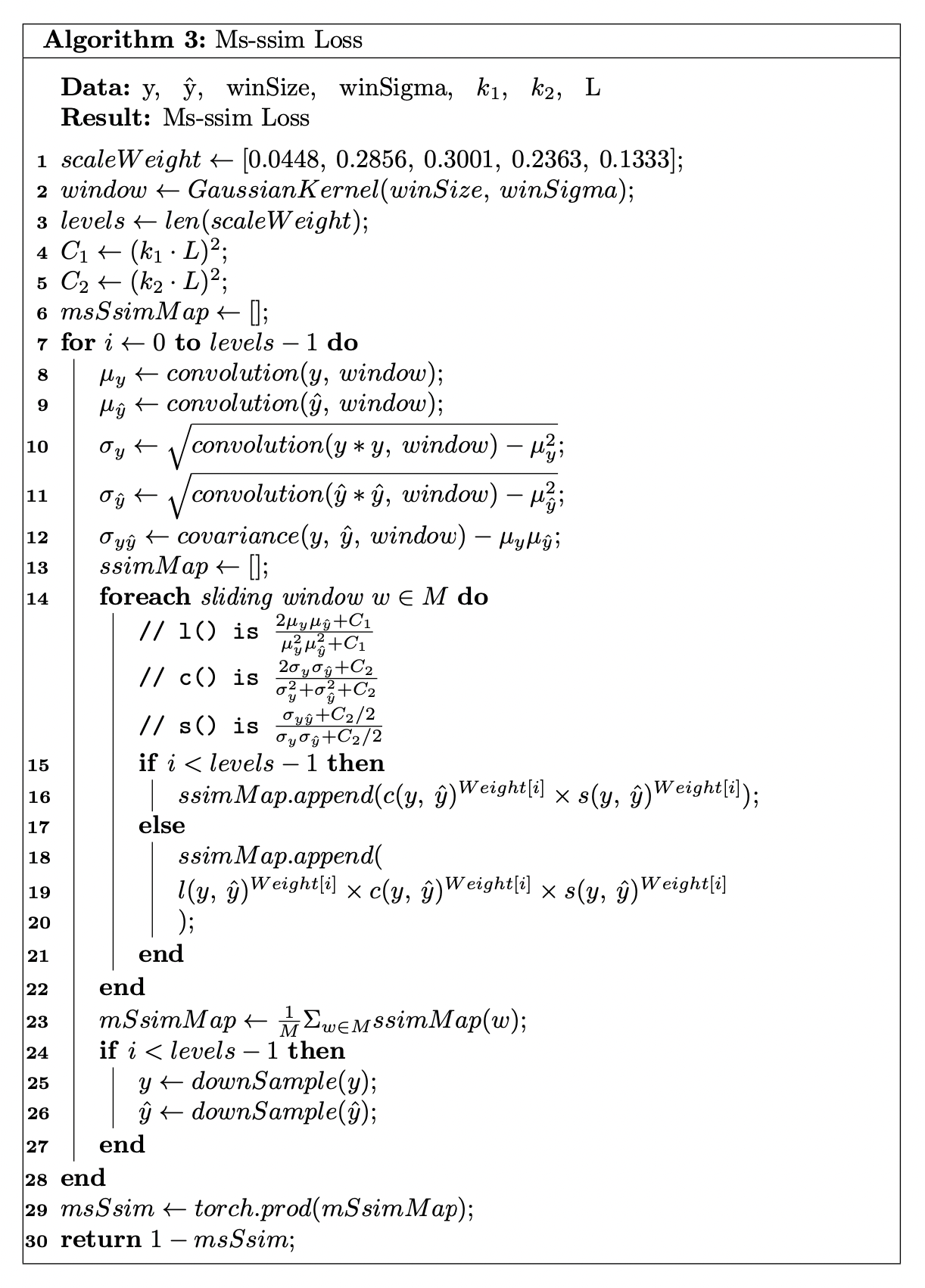


圖13 Ms-ssim Loss 偽代碼

### DSC應用於胎兒頭部分割

本論文用來評估我們預測的圖像與超音波師的標注的重疊率，達到分割層級的評估，公式是：

|  |  |
| --- | --- |
|  | (37) |

：我們預測的圖像的每個像素是胎兒頭部的機率，。

：為超音波師標注的每個像素是胎兒頭部的機率，。

最後因為我們要loss function越低越好，所以用1減去(3‑5)得：

|  |  |
| --- | --- |
|  | (38) |

圖14為我們在此使用的的偽代碼，超參數設置為以下：

Smooth = 0.00005 // 常數，防止分母近似0造成不穩定。

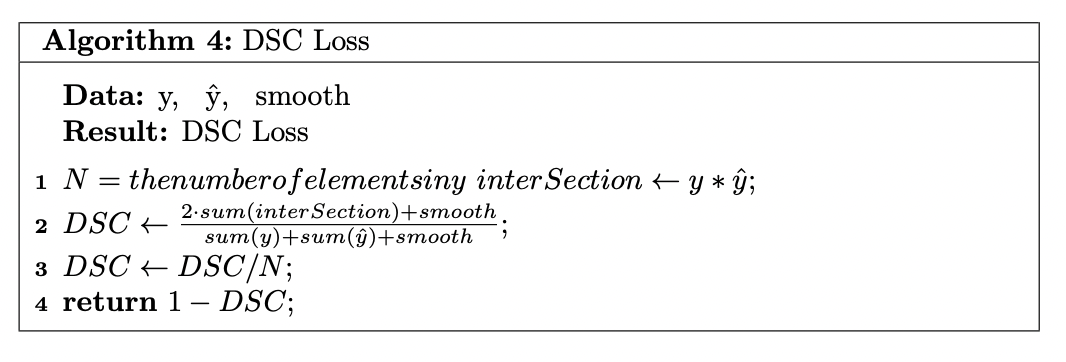


圖14 DSC Loss偽代碼

### 提出的混合損失函式

結合上面(34)(36)(38)得到我們的混合loss function(39)：

|  |  |
| --- | --- |
|  | (39) |

圖15為我們利用混合Loss Function 來訓練模型的權重的偽代碼。

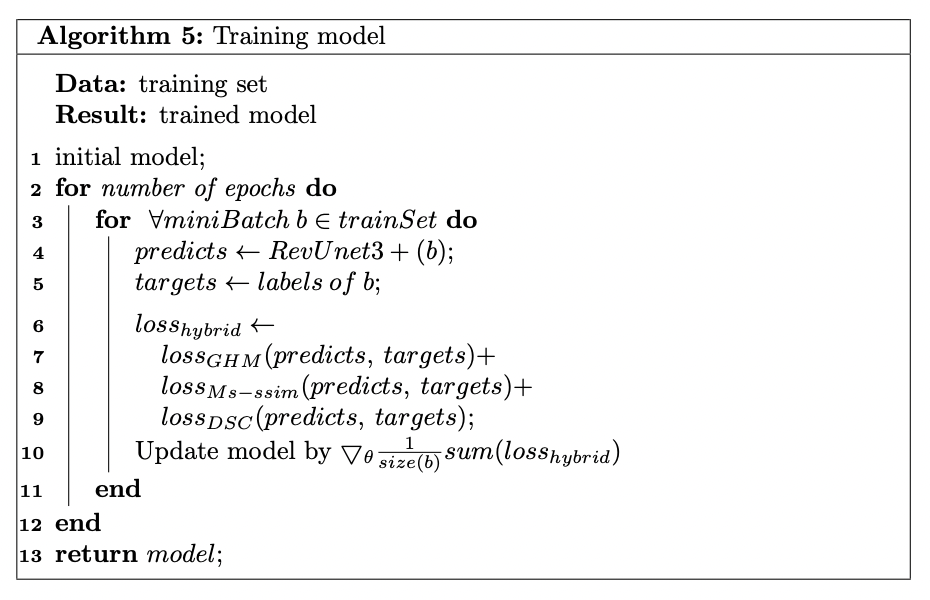


圖15訓練模型偽代碼

## 計算HC值

我們將RevUnet3+ 輸出的二值化的分割圖後，使用OpenCV的findContours() 找尋分割輪廓點，並取最多點的輪廓當作後續橢圓擬合的點集合，擬合採用fitEllipse()方法得到描述橢圓的5個參數，最後把半長軸和半短軸乘上該超音波造影的像素大小，並計算橢圓周長，得出真實胎兒的HC值。

# 實驗與結果

## 資料集來源與前處理

實驗資料集來自醫學公開挑戰平台(<https://grand-challenge.org>)的HC18胎兒頭部超音波資料集(<https://hc18.grand-challenge.org>)[29]，有999張訓練資料，335張驗證資料，混合了3個孕期的胎兒頭部造影資料，每張影像尺寸約為800 × 540 pixels，每個pixels大小為[0.052, 0.326] mm，所有的胎兒造影都是發育正常的。

我們對原始資料集中的標注做修改，填滿了橢圓的內部圖16，並刪除其中4張橢圓超出圖像的造影，共995張，這是為防止資料增量時橢圓不完整造成機器學習錯誤，最後將超音波造影縮小4倍輸入神經網路，尺寸為200 × 135 pixels。

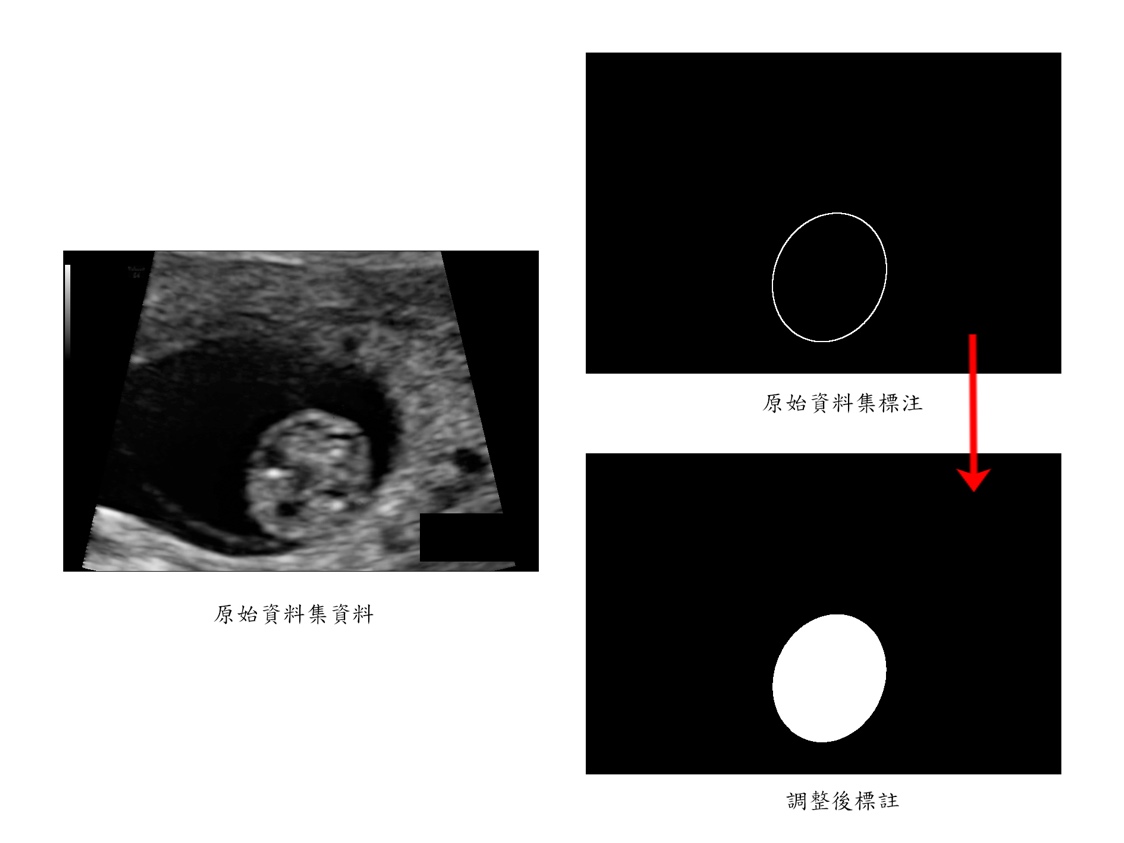


圖16調整原始資料標註

## 資料增量

我們把995張訓練資料刪除分出一些當作測試資料，比例為8：2，訓練資料796張，測試資料199張。然後只對訓練資料做資料增量圖17，將796張訓練資料擴展到23880張，操作為：

1. 旋轉，左右範圍[-25, 25]度，間隔5度。
2. 上下左右反轉。
3. 縮放，[0.85, 1.15]倍，間隔0.05度。
4. 曝光調整，gamma值為[0.5, 0.15]，間隔為0.05。
5. 平移，水平垂直方向範圍[-30, 30]像素，間隔5像素。

對標注不做4變化，其他都有做，並限制不會產生出超出邊界的胎兒頭部造影。

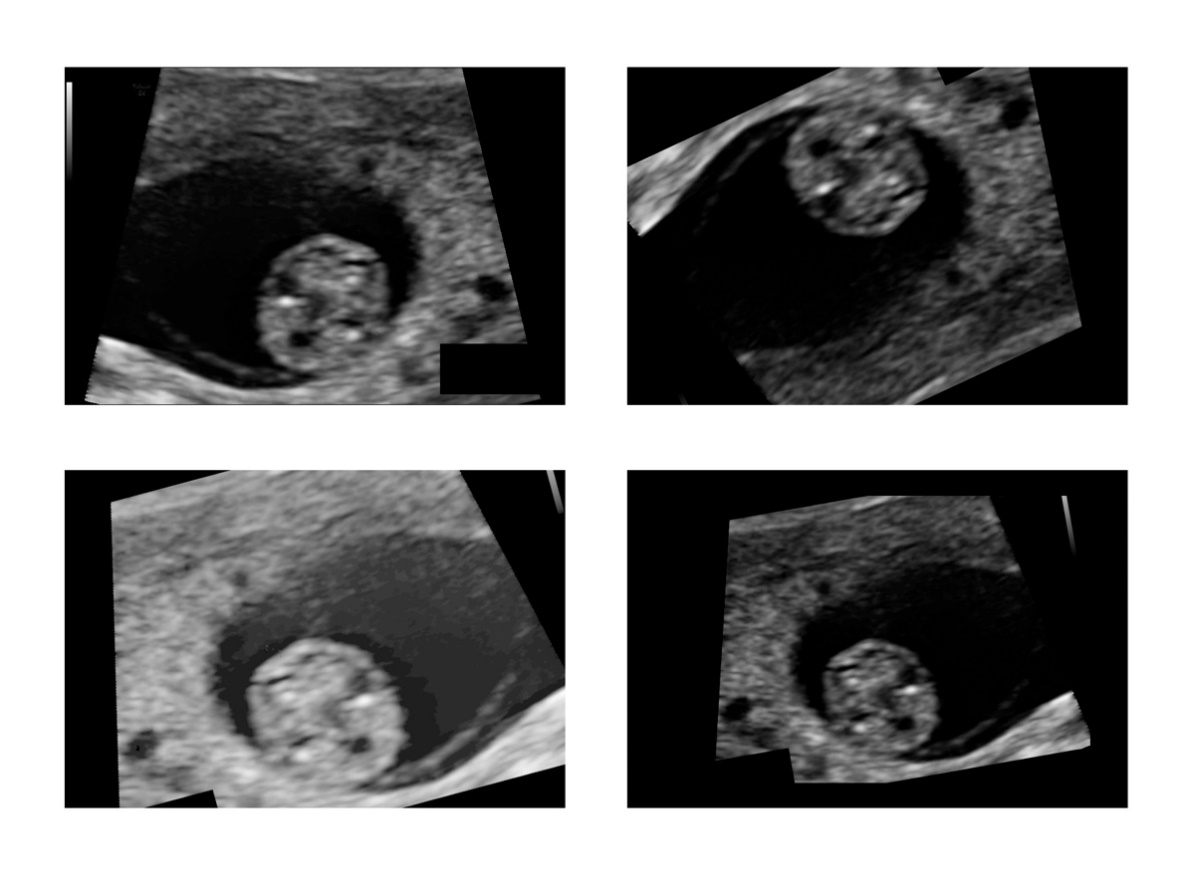


圖17資料增量例子

## RevUnet3+注意力機制可視化

圖18為其中一個樣本的造影及標注解答，圖19為該樣本的預測結果及淺層到深層的注意力機制所關注的地方，越紅則權重越大，越紫色為權重越小，，由上往下為10個epoch的變化，可以發現越淺層的注意力機制注重的是輪廓，而深層的則是位置。



圖18測試注意力機制的原始資料

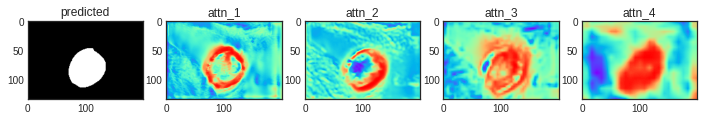
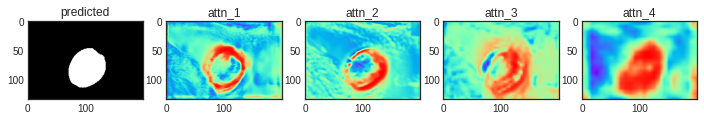
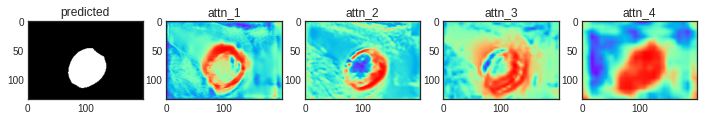
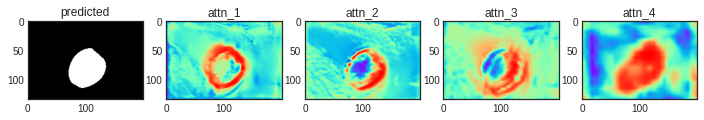
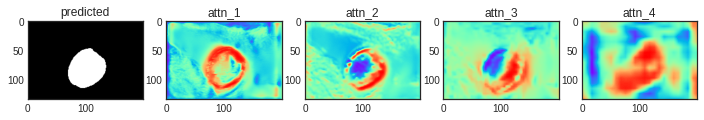
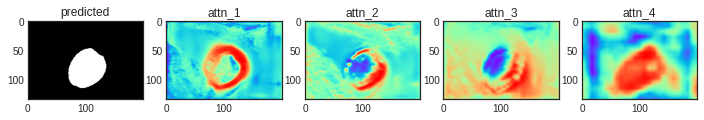
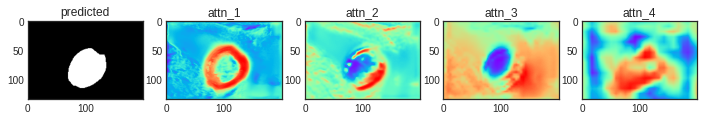
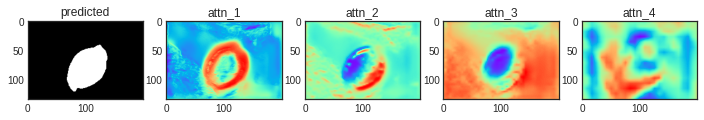
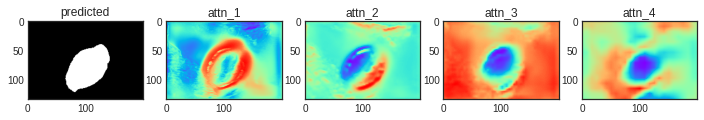


圖19各層注意力機制所關注的地方，由上往下10個epochs的變化

## 評估指標介紹

本論文使用4種指標來判斷系統的表現，分別是：DSC(Dice coefficient)、HD(Hausdorff Distance)、DF (Difference)和ADF (Absolute Difference)。

### DSC(Dice Coefficient)指標

和我們混合loss function裡的DSC一樣，評估預測的圖像與超音波師的標注的重疊率。公式為：

|  |  |
| --- | --- |
|  | (40) |

：為我們預測的圖像。

：為超音波師的標注圖像。

### HD(Hausdorff distance)指標

評估預測的胎兒頭部輪廓像素與超音波師的標注的胎兒頭部輪廓像素之間的距離，外形相似程度，分別從兩個方向看，計算從預測輪廓到標註的距離和標注到預測輪廓的距離。

|  |  |
| --- | --- |
|  | (41) |

|  |  |
| --- | --- |
|  | (42) |

然後比較(41)和(42)取最大者就是HD。

|  |  |
| --- | --- |
|  | (43) |

：取距離。

：為我們預測的圖像。

：為超音波師的標注。

：為我們預測的圖像像素集合中的其中一個像素的座標。

：為超音波師的標注的像素集合中的其中一個像素的座標。

### DF(Difference)指標

評估我們預測的HC週長和超音波師預測的HC，有沒有出現偏移。

|  |  |
| --- | --- |
|  | (44) |

：為我們預測的胎兒頭部周長。

：為超音波師測量的胎兒頭部周長。

### ADF(Absolute Difference)指標

評估我們預測的HC週長和超音波師預測的HC的誤差。

|  |  |
| --- | --- |
|  | (45) |

：為我們預測的胎兒頭部周長。

：為超音波師測量的胎兒頭部周長。

## 實驗結果比較

### RevUnet3+ 三個期孕期指標分數

胎齡醫學上會分三個孕期觀察，每三個月一期，共三期，表1為表示我們的網路在每一期的各項指標結果。

表1 RevUnet3+三個孕期的指標分數

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ADF(mm) | DSC(%) | HD(mm) | DF(mm) |
| 1st trimester | 1.33 ± 1.40 | 96.76 ± 1.81 | 0.89 ± 0.51 | 0.47 ± 1.88 |
| 2nd trimester | 1.58 ± 1.33 | 98.10 ± 0.81 | 1.18 ± 0.57 | 0.02 ± 2.07 |
| 3rd trimester | 2.80 ± 2.39 | 98.08 ± 0.80 | 2.01 ± 0.80 | -0.38 ± 3.69 |

### HC Grand Challenge 排名及平均比較結果

在表2可以看到我們在ADF獲得第一，ADF為此挑戰賽裡面權重最大的指標，因為它直接表示和胎兒HC的誤差，但在DSC及HD還有進步的空間。

表2 RevUnet3+與挑戰賽比較成果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ADF(mm) | DSC(%) | HD(mm) | DF(mm) |
| 2nd | 1.716 ± 1.55 | **98.10 ± 0.93** | **1.17 ± 0.64** | -0.05 ± 2.31 |
| 3rd | 1.74 ± 1.66 | 98.00 ± 1.07 | 1.22 ± 0.68 | **-0.02 ± 2.41** |
| 4th | 1.76 ± 1.66 | 97.97 ± 1.11 | 1.23 ± 0.68 | -0.24 ± 2.44 |
| 5th | 2.8 ± 3.3 | 98.09 ± 0.95 | 1.19 ± 0.66 | 0.15 ± 2.42 |
| Ours | **1.715 ± 1.60** | 97.88 ± 1.16 | 1.25 ± 0.69 | 0.04 ± 2.35 |

### 經典分割網路與其他研究該HC議題論文提出的網路比較結果

在表3可以看出在其他研究相關議題的論文中，我們的HC誤差值最小，遠高過去其他論文作者，但是在DSC的部分卻沒有達到高精準，這可能是由於我們縮小輸入神經網路的造影而讓一些細節消失，縮小造影是因為硬體上的限制，我們在此使用Google Colab Pro。

表3其他論文及經典分割網路和我們的比較結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ADF(mm) | DSC(%) | HD(mm) | DF(mm) |
| Unet | 1.79 ± 1.69 | 97.78 ± 1.17 | 1.31 ± 0.72 | -0.69 ± 2.37 |
| Unet++ | 1.718 ± 1.61 | 97.89 ± 1.12 | 1.29 ± 0.72 | -0.30 ± 2.34 |
| Unet3+ | 1.81 ± 1.66 | 97.97 ± 1.11 | 1.23 ± 0.68 | -0.24 ± 2.44 |
| Heuvel et al.[29] | 2.8 ± 3.3 | 97.0 ± 2.8 | 2.0 ± 1.6 | 0.6 ± 4.3 |
| Zahra et al.[8] | 2.12 ± 1.87 | 96.84 ± 2.89 | 1.72 ±1.39 | 1.13 ± 2.69 |
| Yanyan et al.[9] | 1.84 ± 1.73 | **98.06 ± 1.06** | **1.21 ± 0.69** | - |
| Ours | **1.715 ± 1.60** | 97.88 ± 1.16 | 1.25 ± 0.69 | **0.04 ± 2.35** |

### 與Unet家族視覺化比較分割最差的造影

可以看出在我們提出的網路在經典Unet家族比較中，雖然最後橢圓擬合階段最接近胎兒頭部輪廓，但是在分割階段卻有極小誤分割的部分，圖20中可發現和Unet3+有相同的分類錯誤，推測在多尺度的考量雖然提升了精確分割，但是卻又過於敏感，未來後續可能還需要使用一些其他的濾波器把雜訊給過濾掉。

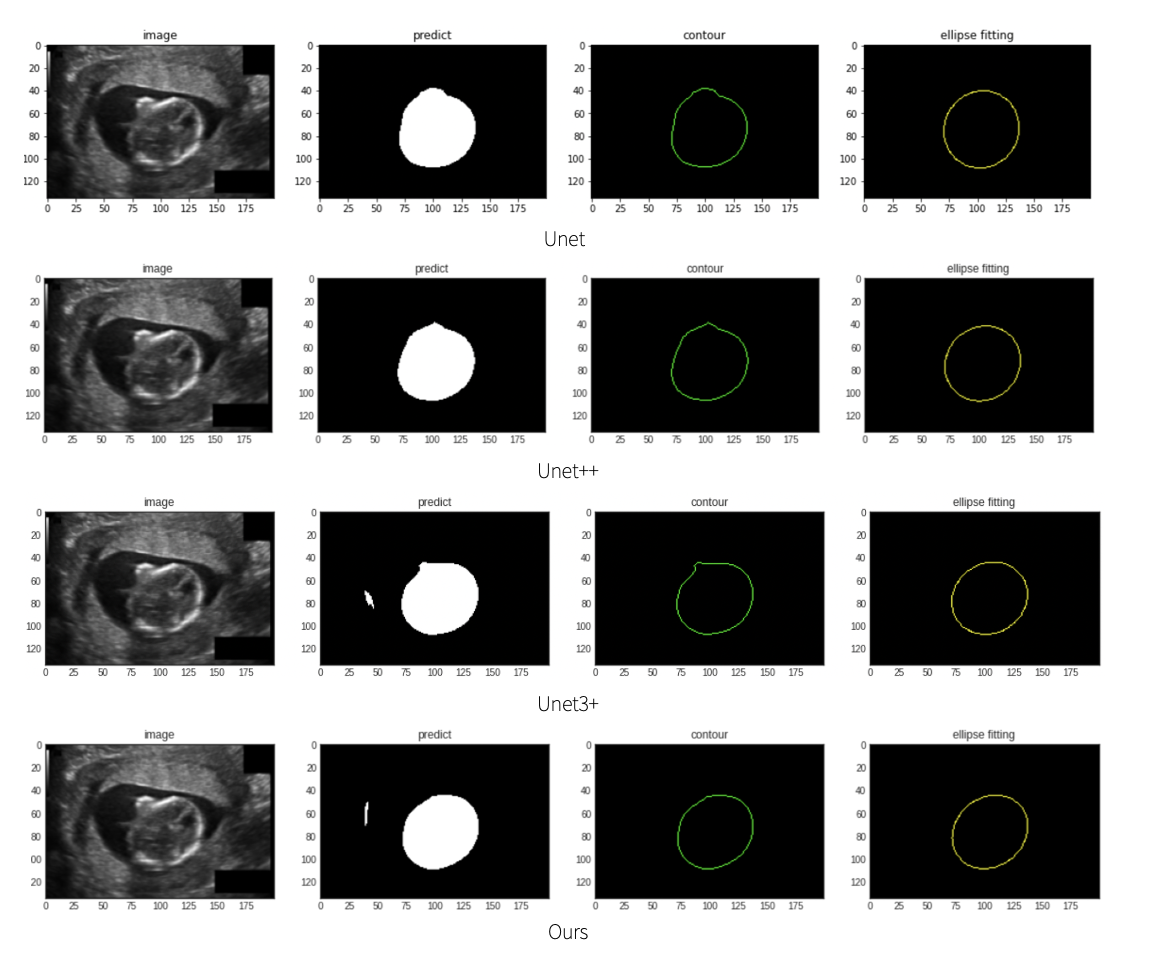


圖20我們提出的網路與Unet家族比較結果

## 訓練超參數設定與執行環境

我們訓練的了40 個世代，學習速率設為0.001，優化器使用Adam，Batch size 設為8，使用DropOut讓節點輸出的特徵圖丟棄了30%的channels，目的為防止過擬合，因為設備限制，圖片縮小1/4，機器學習框架使用Facebook 開發的Pytorch，使用語言是Python，放在Google提供的Colab Pro平台上運行，其所提供的GPU為Tesla P100，表4為其設置表格。

表4超參數設定與執行環境

|  |  |
| --- | --- |
| Epoch | 40 |
| Learning rate | 0.001 |
| Optimization | Adam |
| Batch size | 8 |
| DropOut rate | 30% |
| 圖片縮放 | 倍 |
| 框架 | Pytorch |
| 語言 | Python |
| 平台 | Google Colab Pro |
| GPU | Tesla P100 16GB |

# 結論與未來發展

我們提出的RevUnet3+改良了Unet3+的跳躍連接，認為跳躍連接部分的特徵圖降採樣會和Encoder降採樣損失一樣的細部特徵，所以改採反過來把跳躍連接部分的特徵圖上採樣，這樣也能考慮到多尺度特徵。

目前本研究缺點是只應用於胎兒參數HC的測量，但並未用於其他分割議題上，且雖獲得挑戰賽上第一，但只有ADF指標(挑戰賽排名依據)優於其他論文，其他如DSC、HD未能獲得第一，有可能是因為資料前處理或是橢圓擬合方法等等細微差異導致，但是我想這些微小差異意義並不會太大，充其量只是在過擬合這些訓練資料罷了，若資料有偏頗，如不同造影機器或標註醫師喜好不同等等差異都會影響實際採用的表現，最好的解決方式還是讓訓練資料盡可能多盡可能多樣化，才能保證一定程度上實際使用的表現。

在未來，電腦處理能力變強或超音波圖片變更清晰及資料量變多的話，或許模型的準確率能提高不少，此實驗中，我們使用Google Colab Pro，但此服務提供的GPU有其限制，GPU的RAM太少，所以我們把超音波圖片先縮小4倍，造成精細度消失，這些細節還是很重要的，所以若未來能採用更好的GPU可能採用未縮小的造影圖片來改善系統的準確度。

最後我們期望，不僅只有自動化測量HC來輔助醫生而已，將慢慢擴展顧慮到各種胎兒參數測量並依據這些參數做出擁有自動判別胎兒狀況的全自動化機器。

# 參考資料

1. UNICEF., World Bank., World Health Organization., United Nations. Population Division., and World Health Organization. Reproductive Health and Research., *Trends in maternal mortality, 1990 to 2015 estimates by WHO, UNICEF, UNFPA, World Bank Groupand the United Nations Population Division*. Genève, Switzerland: World Health Organization, 2015.
2. 長庚醫院. “超音波教學與工作規範.” [www.cgmh.org.tw](http://www.cgmh.org.tw). https://www1.cgmh.org.tw/intr/intr5/c6700/OBGYN/Plan/超音波工作規範.htm. (Accessed: 23-Aug-2021).
3. G. Carneiro, B. Georgescu, S. Good, and D. Comaniciu, “Detection and Measurement of Fetal Anatomies from Ultrasound Images using a Constrained Probabilistic Boosting Tree,” *IEEE Transactions on Medical Imaging*, vol. 27, no. 9, pp. 1342–1355, 2008.
4. J. Li, Y. Wang, B. Lei, J. Z. Cheng, J. Qin, T. Wang, S. Li, and D. Ni, “Automatic Fetal Head Circumference Measurement in Ultrasound Using Random Forest and Fast Ellipse Fitting,” *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 215–223, 2018.
5. W. Lu, J. Tan, and R. Floyd, “Automated fetal head detection and measurement in ultrasound images by iterative randomized hough transform,” *Ultrasound in Medicine & Biology*, vol. 31, no. 7, pp. 929–936, 2005.
6. R. V Stebbing and J. E. McManigle, “A boundary fragment model for head segmentation in fetal ultrasound,” *Proc. Chall. US Biometric Meas. from Fetal Ultrasound Images, ISBI*, pp. 9–11, 2012.
7. M. Sinclair, C. F. Baumgartner, J. Matthew, W. Bai, J. C. Martinez, Y. Li, S. Smith, C. L. Knight, B. Kainz, J. Hajnal, A. P. King, and D. Rueckert, “Human-level Performance On Automatic Head Biometrics In Fetal Ultrasound Using Fully Convolutional Neural Networks,” *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018.
8. Z. Sobhaninia, S. Rafiei, A. Emami, N. Karimi, K. Najarian, S. Samavi, and S. M. R. Soroushmehr, “Fetal Ultrasound Image Segmentation for Measuring Biometric Parameters Using Multi-Task Deep Learning,” *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019.
9. Y. Y. Xing, F. Yang, Y. J. Tang, L. Y. Zhang, “Ultrasound fetal head edge detection using fusion UNet++,” *Journal of Image and Graphics*, 25(2), pp. 366-377, 2018.
10. O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *arXiv.org*, 18-May-2015. [Online]. Available: https://arxiv.org/abs/1505.04597v1. [Accessed: 23-Aug-2021].
11. G. E. Hinton, R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
12. Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, “UNet++: A nested U-Net architecture for medical image segmentation,” *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support (2018)*, vol. 11045, pp. 3–11, 2018.
13. H. Huang *et al.*, “UNet 3+: A full-scale connected UNet for medical image segmentation,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020.
14. S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating deep network training by reducing internal covariate shift,” *arXiv [cs.LG]*, 2015.
15. O. Oktay *et al.*, “Attention U-Net: Learning where to look for the pancreas,” *arXiv [cs.CV]*, 2018.
16. P. Anderson *et al.*, “Bottom-up and top-down attention for image captioning and visual question answering,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6077-6086, 2018.
17. D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” *arXiv [cs.CL]*, 2014.
18. P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, “Graph Attention Networks,” *arXiv [stat.ML]*, 2017.
19. F. Wang *et al.*, “Residual Attention Network for Image Classification,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6450-6458, 2017.
20. Y. Ma, Q. Liu, and Z. B. Quan, “Automated image segmentation using improved PCNN model based on cross-entropy,” in *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, pp.* 743-746*,2004*.
21. T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, “Focal loss for dense object detection,” in *2017 IEEE International Conference on Computer Vision (ICCV)*, 42, 318-327, 2017.
22. B. Li, Y. Liu, and X. Wang, “Gradient Harmonized Single-Stage Detector,” in *Proc. Conf. AAAI Artif. Intell.*, vol. 33, pp. 8577–8584, 2019.
23. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.
24. Z. Wang, E. P. Simoncelli, and A. C. Bovik, “Multiscale structural similarity for image quality assessment,” in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, 2004.
25. C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, “Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations,” *Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support (2017)*, vol. 2017, pp. 240–248, 2017.
26. S. Suzuki and K. Be, “Topological structural analysis of digitized binary images by border following,” *Comput. Vis. Graph. Image Process.*, vol. 30, no. 1, pp. 32–46, 1985.
27. A. W. Fitzgibbon and R. B. Fisher, “A buyer’s guide to conic fitting,” in *Procedings of the British Machine Vision Conference 1995*, 1995
28. J. Tompson, R. Goroshin, A. Jain, Y. LeCun, and C. Bregler, “Efficient object localization using Convolutional Networks,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
29. T. L. A. van den Heuvel, D. de Bruijn, C. L. de Korte, and B. van Ginneken, “Automated measurement of fetal head circumference using 2D ultrasound images,” *PLoS One*, vol. 13, no. 8, p. e0200412, 2018.