

# 國立中正大學資訊工程研究所

## 碩士論文

醫學影像專題二

Medical Image Project2

張世亞、蔡旻勳

指導教授：林維暘

中華民國 109 年 5 月

# 誌謝辭

在這門課的兩個月裡，首先要感謝也要對不起我們的指導教授林維暘，我們不是一位非常優秀、十分用功的學生，是老師不厭其煩的一再細心教導與指引方向，令愚笨學生的我們能夠有所成長，才使得本論文得以順利完成，在此謹敬上最高謝意與感激

其次，感謝我們同窗的研究夥伴——軒綸、耘愷和在學期間所有同學、學長姐、學弟妹，有你們與我們共度這兩個月時光讓我們研究生生活更加美好充實與豐富。還要感謝我們大學的好友以及學弟妹們，在這孤單期間有你們陪伴使我們的生活更多彩多姿也讓我們充滿了開心。再來感謝在嘉義的好朋友與好哥們，雖然這段期間並不常見面，但是有你們的關心與打氣，讓我們更有動力能在繁忙的課業中順利撐過來。

另外特別感謝幾位我最重要的貴人——昀倫、宏瑜、佑安、麥克，當我們心情陷入低潮時是你們給予我們支持與鼓勵，讓我們學會成長、讓我們充滿勇氣、讓我們更堅強更勇敢的去面對一切困境，沒有你們也就沒有今天堅強的我們，謝謝你們，你們是我們一生中永遠最棒的好朋友。

最後，我們要感謝我們的父母、家人以及我們最敬愛的上帝，有你們細心栽培與照顧，讓我們無後顧之憂的學習與完成學業，所以我們的成就也是你們的驕傲。謹以此篇論文獻給我們的師長、家人、同學以及所有關愛我們的朋友，謝謝你們。

# 摘要

本篇論文目的在於使用深度學習架構對眼睛醫學影像進行患部偵測並以Mask的方式呈現出患部的範圍及形狀，最終以ROC之曲線下面積AUC評估影像切割的成效。此次的眼部病變資料有40組，每組為同一位病患的影像資料，包括5個時間點的5張影像資料，因此整個資料集總共包括200張眼部病變圖像。從視覺上來觀察，5個時間點的影像差異主要差異在於對比度等細部的差異，在患部形狀和大小上的差異不大。

針對醫學影像的患部辨識問題上，傳統方法大都是採以人工方式進行判斷並框選出患部，而在此次的醫學影像切割問題，我們採用深度學習的方式取代傳統人工辨識，希望能藉此改善人工辨識所導致的誤差及主觀判斷。我們嘗試了5個深度學習架構，分別為 Unet(助教提供)、FCN-ResNet101[1]、SegNet[2]、DeepLabv3及DeepLabv3+[3]，並藉由調整部分超參數進行實驗比較。最終我們利用FCN-ResNet101訓練150個epochs得到表現最優異的模型，最佳AUC\_ROC可達到0.9681。

**關鍵詞：**深度學習、AUC、醫學影像、影像切割、FCN-ResNet101。

## Abstract

The purpose of this paper is to use deep learning framework to detect the affected region of the eye medical image and present the range and shape of the affected part in Mask way. Finally, the area under the curve of the ROC(receiver operating characteristic curve) which is call AUC(Area under curve) is used to evaluate the effectiveness of image cutting. There are 40 groups of eye lesion data this time, each group is the image data of the same patient, including 5 image data at 5 time points, so the entire data set includes a total of 200 eye disease images. From a visual point of view, the main difference in the image difference at five time points is the difference in contrast and other details, and the difference in shape and size of the affected part is not large.

Regarding the identification of affected parts of medical images, traditional methods mostly use artificial methods to judge and select the affected parts. In this medical image cutting problem, we use deep learning to replace traditional manual identification, hoping to improve Errors and subjective judgments caused by manual identification. We tried five deep learning architectures, namely Unet (provided by the teaching assistant), FCN-ResNet101, SegNet, DeepLabv3 and DeepLabv3 +, and experimented by adjusting some hyper parameters. Finally, we used FCN-ResNet101 to train 150 epochs to get the best performing model, and the best AUC\_ROC can reach 0.9681.

**Keywords :** deep learning, AUC(Area under curve), medical images, Image Segmentation, FCN-ResNet101.

# 目錄

介紹.....	5
方法.....	5
實驗.....	7
結論.....	12
參考.....	12

# 一、介紹

在當前的醫療診斷中，醫學圖像具有越來越重要的臨床應用價值。常用的醫學圖像有磁共振（Magnetic Resonance Imaging，MRI）、超聲、電腦斷層掃描（X-Ray Computer Tomography，CT）、正電子發射電腦斷層掃描以及病理切片圖像等[1]。傳統醫學影像處理及判斷大多仰賴專家或是醫師以個人經驗進行判斷，但這樣的方式容易受到個人主觀判斷的影響而造成誤差。此外，隨著影像資料的急遽增加，傳統人工方式難以應付大量的資料並做出客觀分析。隨著硬體運算能力的提升及深度學習模型的提出，利用GPU運算搭配深度學習演算法解決醫學影像分析變得越來越常見。

而我們本篇論文目的便是使用深度學習進行醫學影像的患部影像切割。在本次實驗中我們分別採用了 Unet(助教提供)、FCN-ResNet101、SegNet、DeepLabv3及DeepLabv3+ 五種架構加上超參數調整並搭配不同資料集切割方式進行試驗，試圖找出最佳的影像切割模型以精準切割出眼睛患部的形狀集範圍，最終表現最佳的模型為 FCN-ResNet101訓練150個epochs，其最佳AUC的結果可達0.9681。

# 三、方法

在本次的醫學影像切割問題上，我們實驗了5個不同的模型，並針對以下幾個方向進行調整，試圖找出最佳的模型集實驗數據組合。調整方向包括以下幾個小節：3.1 dataset 劃分，3.2 參數調整，3.3 模型更換。

## 3.1 dataset 劃分

為了測試各模型並找出模型的最佳參數組合，我們的資料集切割分成兩種方式：

1. 模型測試：所有模型的 baseline 都是將訓練集與驗證集採 9:1 的方式劃分。
2. 交叉驗證：確認模型在測試階段為可行後，便會進行 5 次交叉驗證，故訓練集與驗證集的劃分比為 4:1，會產生五個模型，取最佳的模型當作最後的結果。

## 3.2 參數調整 我們的參數調整分成以下幾種：

### Epochs：

針對不同模型，我們嘗試了 10 ~ 150 個 epochs，其中 Unet 在 50 個 epochs 已經大致收斂；FCN-ResNet101 在 150 個 epochs 的表現佳；SegNet 約在 150 個 epochs 的表現佳；DeepLabv3 在 50 個 epoch 的視覺效果比 70 或 100 個 epochs 的表現還要好；DeepLabv3+ 在 100 個 epochs 的表現佳。

### Learning rate :

我們的 learning rate 有經過以下調整 :  $0.001 \rightarrow 0.0002 \rightarrow 0.001$  進行測試, 最終結果為 0.001 表現最佳。

### Batch size :

有做些微的調整 :  $1 \rightarrow 2 \rightarrow 1$ , 但當設為 2 時, 某些模型訓練會發生梯度爆炸或梯度消失的情況, 導致 NaN 發生, 故最終以 1 當作穩定的 batch size。

### Optimizer :

針對不同的模型, 我們嘗試了兩種 optimizer, 包括 RMSprop (Unet) 及 SGD (FCN-ResNet101、SegNet、DeepLabv3、DeepLabv3+)。

### Loss function :

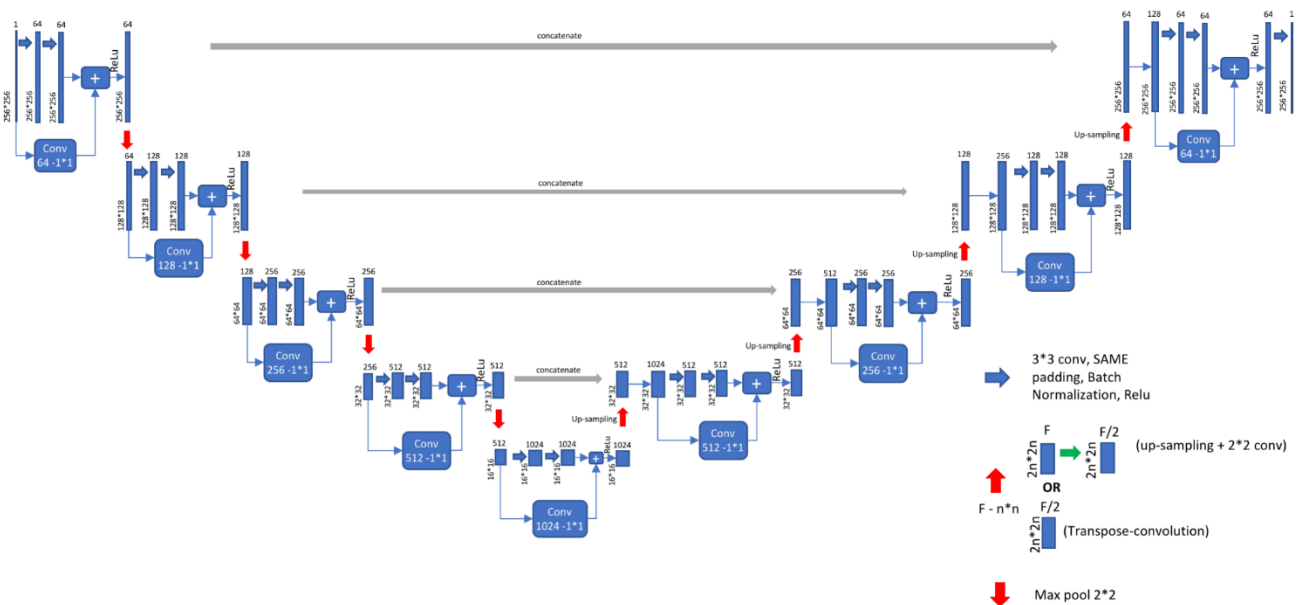
針對不同模型, 我們嘗試了兩種 loss function, 包括 BCEWithLogitsLoss (Unet) 及 softmax loss (FCN-ResNet101、SegNet、DeepLabv3、DeepLabv3+)。

## 3.3 模型更換 我們的 model 採用以下幾種

### I. Unet

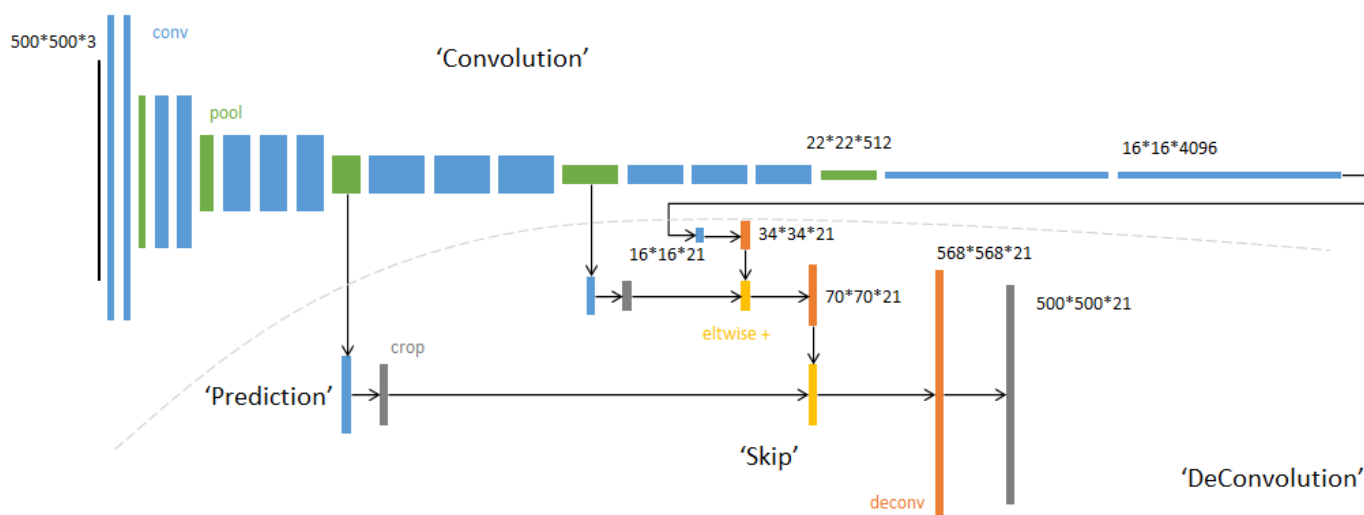
整體結構主要分為 2 部分, 包括左半部的 contraction 及右半部的 expansion, 如下圖所示。Contraction 的部分是由 4 個 down sample 的模組所組成, 每個模組包括兩個  $3 \times 3$  的 convolution 和一個  $2 \times 2$  的 max pooling; expansion 則是由 4 個 up sampling 所組成, 每做一次 up sampling 便會與前面 down sample 階段得到的 feature (經過 crop 後) 進行拼接, 最終得到跟原圖一樣大小的輸出結果。

Unet 常被應用在各領域的影像切割問題, 在生物醫學圖像上的表現尤佳, 對於有限的圖像資料集也能有相當不錯的表現。



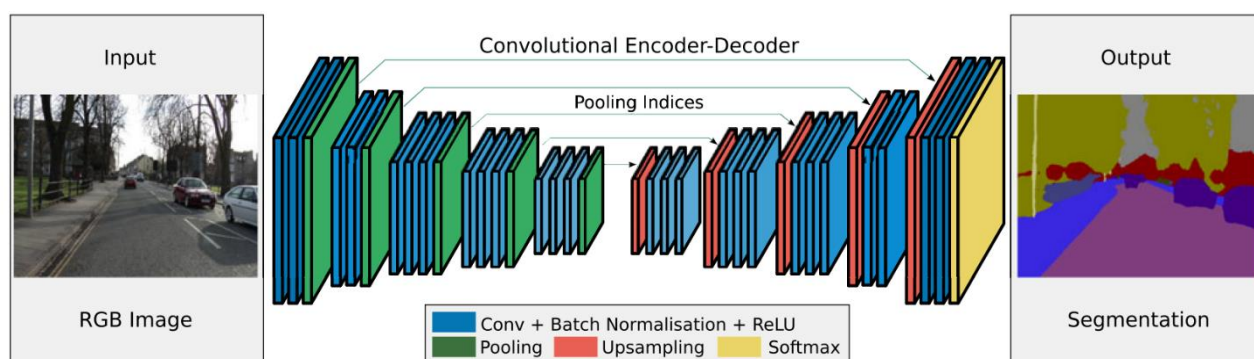
## II. FCN-ResNet101

下圖為FCN-ResNet101的架構圖，虛線上半部為convolution和pooling所組成的全卷積層，負責特徵擷取；而虛線下半部則針對某些convolution階段進行預測及Up sampling。其Fully convolutional的架構能幫助模型在取feature的過程中維持像素與像素之間的空間意義。



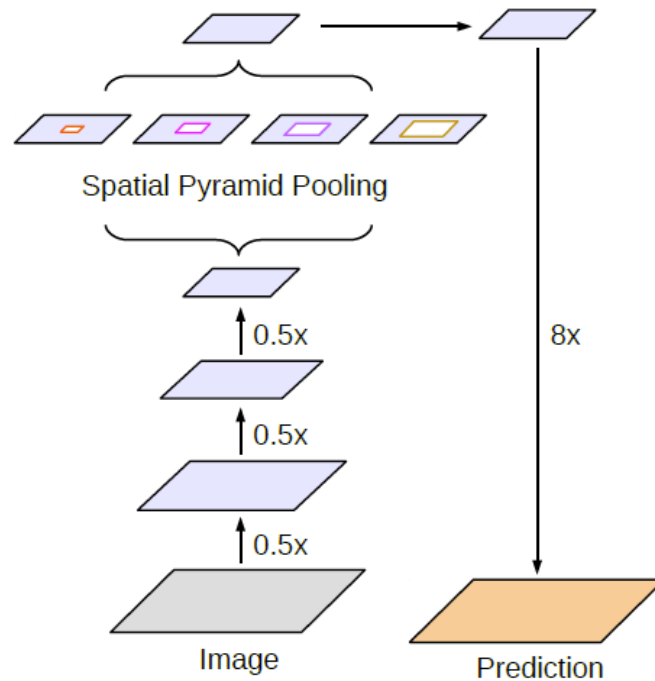
## III. SegNet

下圖為SegNet架構的示意圖。主要是由卷積層所組成，沒有全連接的層。解碼器使用來自其編碼器的傳輸池索引對輸入進行上取樣，以生成稀疏特徵圖。此外，它可與訓練的濾鏡庫進行卷積以濃縮特徵圖。最終的解碼器輸出特徵圖被饋送到soft-max分類器，以進行像素分類。



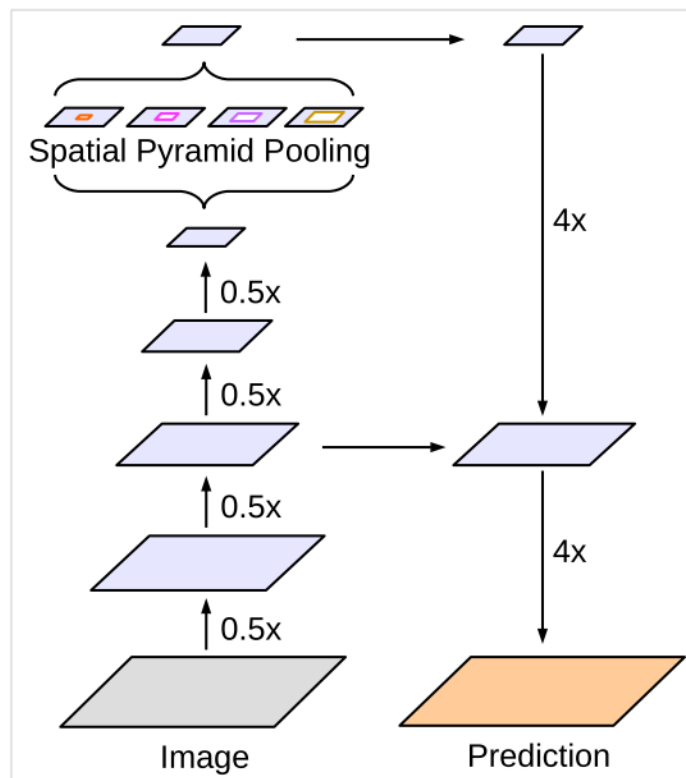
## IV. DeepLabv3

下圖為DeepLabv3的架構，其Spatial Pyramid Pooling的設計讓模型能夠得到不同scale的feature，因此比起deeplabv2，此架構能夠捕捉多種scale的物件訊息。進而再多物件分割問題上有更好的表現。



## V. DeepLabv3+

下圖為DeepLabv3+的架構，它利用DeepLabv3當作encoder，並在此基礎上再加上一個decoder對分割結果做更細微的處理。此外DeepLabv3+新增了Xception結構應用於分割任務中，在 ASPP 和解碼器模塊中加入深度分離卷積，獲得到強大又快速的模型。





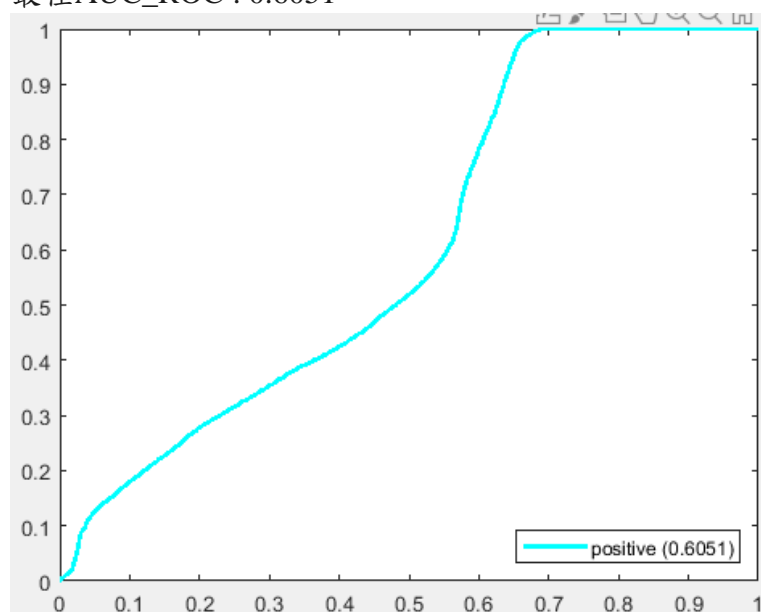
## 四、實驗

針對3.2所條列出來的調整方案，我們在4.1說明調參細節和各模型的最佳AUC，總共嘗試了4個實驗組並與1個default設定的對照組進行比較。在4.2的部分我們將比較各模型的成效及優劣並展示最佳模型：FCN-ResNet101的40張AUC圖。

### 4.1 實驗實作細節

#### 4.1.1 Unet

- 說明：原始的設定是遵照老師所給的範例程式及參數，其中在epoch的部分我們有進行2次調整。但從結果來看，原始設定的模型在50個epoch之後就沒有太多的變化。
- 參數設定
  - Model : Unet
  - Epoch : 10 → 50 → 100
  - Learning rate : 0.001
  - Loss function : BCEWithLogitsLoss
  - Optimizer : RMSprop
  - Batch size : 1
- 結果
  - 最佳AUC\_ROC : 0.6051

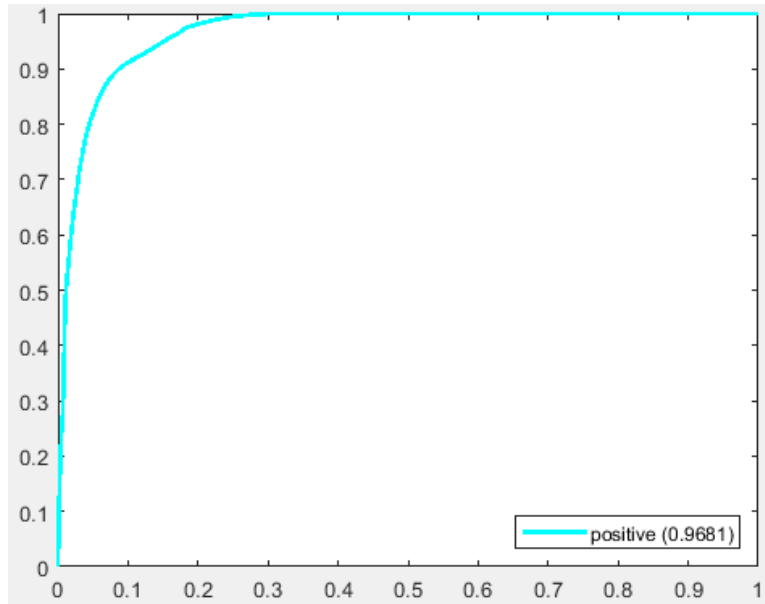


#### 4.1.2 FCN-ResNet101

- 說明：整體效果隨著epoch增加有所提升，當epoch提升至150時Mask效果比其他模型還要優異。而當batch size調整為2時，有時會發生梯度爆炸或梯度消失的情形，造成訓練無效，故最終調整為1。

- 參數設定
  - Model : FCN-ResNet101
  - Epoch : 100 → 70 → 120 → 150
  - Learning rate : 0.001
  - Loss function : 各pixel取softmax loss
  - Optimizer : SGD
  - Batch size : 1 → 2 → 1

- 結果
  - 最佳AUC\_ROC : 0.9681

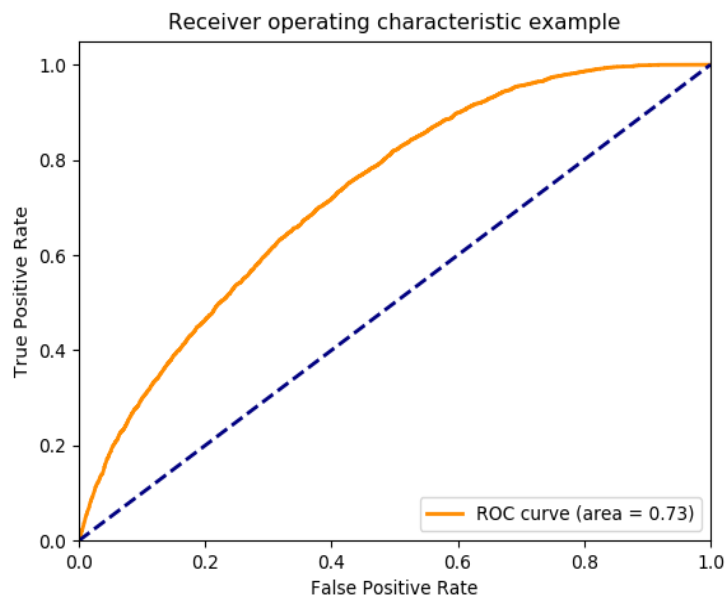


#### 4.1.3 SegNet

- 說明：這個模型根據5個時間點採用交叉驗證方式對資料集進行劃分，總共訓練了5個模型。嘗試調整Epoch從100提升至150，發現效果並無顯著提升。而初始學習率則由 0.001 調整至 0.0002，達到最佳效果。最終實驗結果以 1、3、4、5 時間點為訓練集，2 時間點為驗證集所得到的模型結果為最佳。模型整體表現穩定，但效果不佳。

- 參數設定
  - Model : SegNet
  - Epoch : 100 → 150
  - Learning rate : 0.0002 → 0.001 → 0.0002
  - Loss function : 各pixel取softmax loss
  - Optimizer : SGD
  - Batch size : 5

- 結果
  - 最佳AUC\_ROC : 0.73

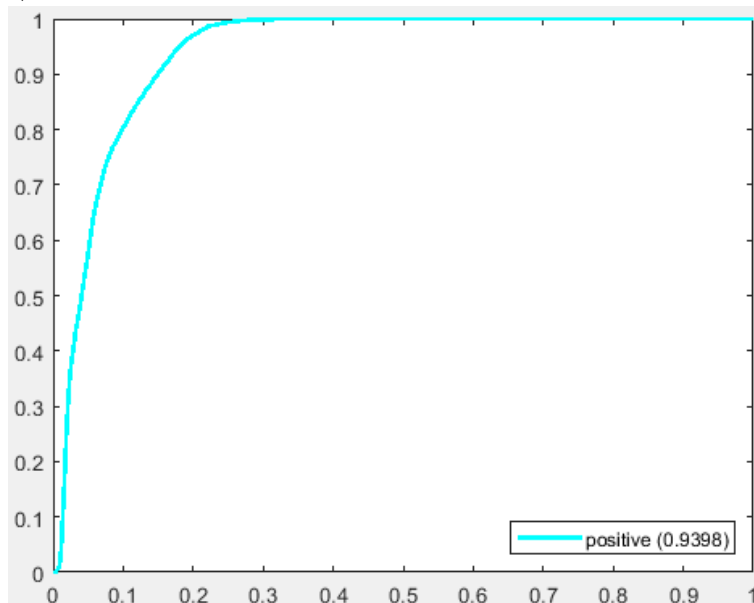


#### 4.1.4 DeepLabv3

- 說明：DeepLabv3在50個epochs時的視覺效果最佳，而在batch size的調整上與FCN-ResNet101發生了一樣的問題。另外在資料集劃分的部分，當我們將訓練集:驗證集設為0.55:0.45時，模型的表現比0.9:0.1的劃分方式還要佳。

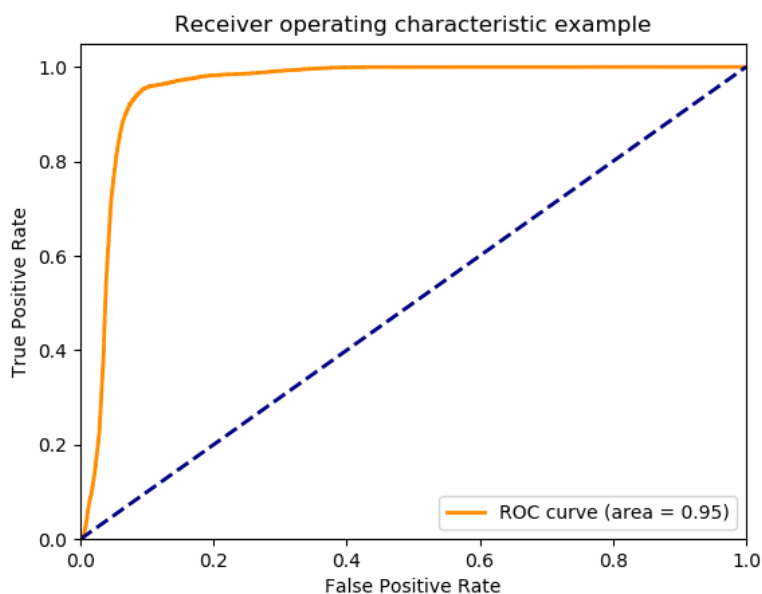
- 參數設定
  - Model : DeepLabv3
  - Epoch : 50 → 100 → 70
  - Learning rate : 0.001
  - Loss function : 各pixel取softmax loss
  - Optimizer : SGD

- 結果
  - 最佳AUC\_ROC : 0.9398



#### 4.1.5 DeepLabv3+

- 說明：這個模型根據5個時間點採用交叉驗證方式對資料集進行劃分，總共訓練了5個模型。DeepLabv3+ 相對於 DeepLabv3 有了小幅度的效能提升，但缺點同樣是表現不夠穩定，對資料的預測大起大落。另外本次實驗模型的 backbone 同樣嘗試了 resnet、drn 以及 Deeplabv3+ 新增的 xception 三種架構，其中表現最佳，不穩定情況最低的架構為 resnet，所以本次Deeplabv3+的實驗結果採用resnet作為代表，搭配時間點進行訓練，共獲得總計15個模型。
- 參數設定
  - Model : DeepLabv3+
  - Epoch : 100 → 150
  - Learning rate : 0.0002 → 0.001
  - Loss function : 各pixel取softmax loss
  - Optimizer : SGD
  - Batch size : 5
- 結果
  - 最佳AUC\_ROC : 0.95



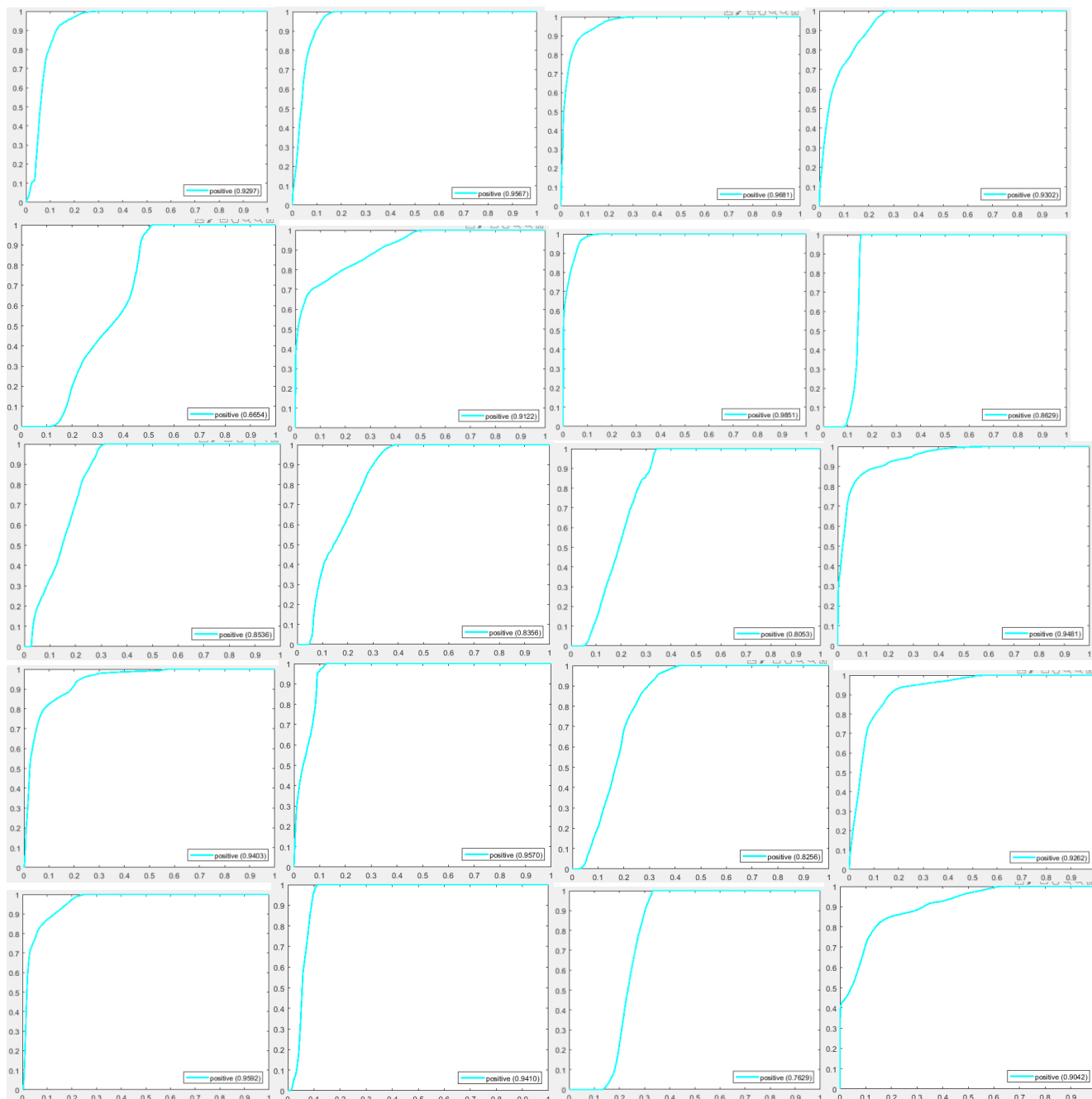
## 4.2 比較

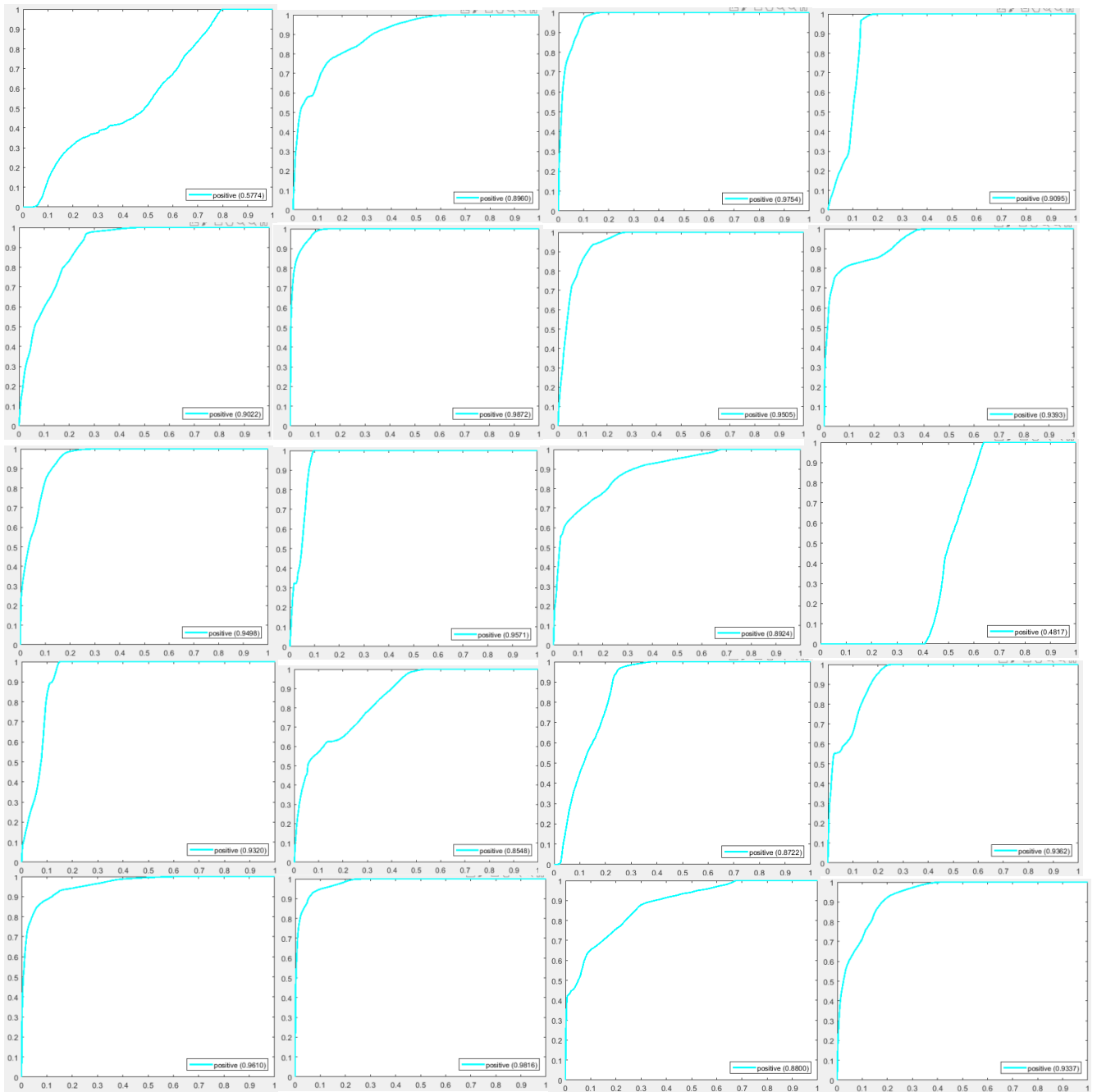
### 4.2.1 比較總表

模型	最佳AUC	評估結果	排名
Unet	0.6051	Baseline	5
FCN-ResNet101	0.9681	最佳	1
SegNet	0.73	故圖片表現穩定，但效果有限，平均AUC無法提高太多	4
DeepLabv3	0.9398	表現不穩定，少部分圖片的	3

		AUC可能低於0.5。	
DeepLabv3+	0.95	表現不穩定，少部分圖片的 AUC可能低於0.5。	2

#### 4.2.2 最佳模型FCN-ResNet101之40張AUC結果圖





## 五、結論

本次的醫學影像專題中主要是針對視網膜中兩種病理現象進行辨識，由於訓練資料不算充足，所以適合使用參數量較小的模型進行訓練，其中 ResNet 的 residual block 使得它在這次的專題中能有相當優異的表現。而在資料分布的部分，這次的資料有相同病人的圖片聚集在一起的現象，導致訓練和測試資料的切分方式也會影響這次的實驗結果。

總體而言，我們這次的貢獻大致有下列幾項：

1. 觀察圖片在資料夾中的分布，並提供不同劃分 dataset 的方式，最後加以驗證。
2. 分析不同模型的優劣，選擇出適當的模型，並將準確率提升到最高 98.6111%
3. 針對不同的參數進行調整，並找出穩定且表現優異的參數組合。

就本次實驗的設計規劃方面來說，對模型以外的改善方面仍有提升空間。針對本次實驗不足之處，我們預計採用其他方案搭配我們的模型實驗方案來提升效能。預計採用的方案如下：

1. 針對資料不足的部分採用資料擴增的方式增加資料量。
2. 針對 optimizer 進行優化，採用更合適的優化器方案。
3. 嘗試改進 loss function，從 BCELoss 改進成 CrossEntropy loss 等等。
4. 針對學習率部分，嘗試移除動態學習率，確認是否有改善畫面全黑的情形。

## 六、參考

- [1] Jonathan Long\*, Evan Shelhamer\*, Trevor DarrellFully. Convolutional Networks for Semantic Segmentation. In Computer Vision and Pattern Recognition, 2015.
- [2] Vijay Badrinarayanan, Alex Kendall, Roberto Cipolla, Senior Member. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. In Computer Vision and Pattern Recognition, 2016.
- [3] Liang-Chieh Chen George Papandreou Florian Schroff Hartwig Adam Google Inc. Rethinking Atrous Convolution for Semantic Image Segmentation. arXiv:1706.05587.
- [4] Goel N,Yadav A,Singh BM.Medical image processing: A review[A].Computational Intelligence on Power, Energy and Controls with Their Impact on Humanity[C].2017:57-62.