

國立雲林科技大學資訊管理系

機器學習-作業二

Department of Information Management

National Yunlin University of Science & Technology

Assignment

Endotracheal Tube 氣管內管資料集

Endotracheal Tube Endotracheal Tube DataSet

巫宇哲、鄭皓名、翁振洋

指導老師：許中川 博士

Advisor: Chung-Chian Hsu, Ph.D.

中華民國 112 年 5 月

May 2023

## 摘要

本研究透過 ETT 資料集中的氣管 X-ray 影像切割圖來進行模型訓練，以 CNN 卷積神經網路為模型架構，並運用深度學習技術來做這項研究，主要是採用 AI 影像切割的技術，本研究主要的動機是想要幫助院方在執行插管任務的時候，不會因為插管插太深導致內臟破裂與插管插太淺運送不足夠的氧氣，能幫助醫生用最準確的方式將氧氣管順利的放入正確的位置。而目的是訓練和驗證真實的氣管遮罩圖與本研究預測的氣管遮罩圖，之間是否存在差異，並觀察當中的誤差值是多少，在從誤差值當中來對模型進行修改與調整，主要是調整模型當中的一些參數，來達到更準確的預測，有準確的插管資訊，醫生也能降低插管失敗的風險，並希望能藉由提供正確插管資訊，來幫助醫療人員成功執行插管任務。本研究使用已經切割成訓練、驗證與測試集的氣管 X-ray 影像資料，並且各資料集皆已標註好氣管內管插管的遮罩(mask)，希望能藉由與影像切割相關的深度學習網路，例如：TransUnet、UNet++等，達到順利切割氣管內管之目標。最後這個實驗我們使用了 transunet 的架構，他是結合了 unet 和 transformer 兩種模型下去組合的，使用 ReduceLROnPlateau、EarlyStopping 等 callbacks 來動態調整超參數，我們最後所用的 set\_4 參數組合參數量高達 65,991,809，那訓練的結果我們可以看到 Fold1-Fold5 其實績效都是差不多的，只有 Fold3 相較起來差一點，我們也嘗試調了不同種的參數組合，但其實都影響不大，最後我們選擇 set\_4 作為後續測試，因為它 IoU 指標相對最好 0.4707，那它主要的改變是損失函數的計算，另外我們發現動態的學習率調整也可能讓模型有機會學得更好，從上面實驗結果可以發現有些模型適當的調整學習率，是可以讓驗證集突破現有的績效的。

關鍵字：TransUnet、UNet、Image segmentation、ReduceLROnPlateau、EarlyStopping

# 一、緒論

## 1.1 動機

本研究透過 ETT 資料集中的氣管 X-ray 影像切割圖來進行模型訓練，以 CNN 卷積神經網路為模型架構，會做這項研究主要動機是想要幫助院方在執行插管任務的時候，不會因為插管插太深導致內臟破裂與插管插太淺運送不足夠的氧氣，能幫助醫生用最準確的方式將氧氣管順利的放入正確的位置，這件事情是相當重要的，準確的插管病人才能夠有充足的氧氣，插管主要的功能是輔助病患呼吸，確實無誤的插管才是對病患最有益處的。

## 1.2 目的

本研究想要藉由深度學習的方式，還有運用 AI 影像切割的技術，來訓練和驗證真實的氣管遮罩圖與本研究預測的氣管遮罩圖，之間是否存在差異，並觀察當中的誤差值是多少，在從誤差值當中來對模型進行修改與調整，主要是調整模型當中的一些參數，來達到更準確的預測，有準確的插管資訊，醫生也能降低插管失敗的風險，插管是一項困難的任務，本研究希望能藉由提供正確插管資訊，來幫助醫療人員成功執行插管任務。

# 二、方法

## 2.1 實作說明

我們使用已經切割成訓練、驗證與測試集的氣管 X-ray 影像資料，並且各資料集皆已標註好氣管內管插管的遮罩(mask)，希望能藉由與影像切割相關的深度學習網路，例如：TransUnet、UNet++等，達到順利切割氣管內管之目標。在丟入模型訓練前，我們有先做一系列的資料前處理，如：將 X-ray 影像切割到適合模型學習的大小、切割後對影像 resize、利用資料增強的方式對影像平移及翻轉，最後針對每個已劃分好的 mask 做二值化處理，主要目的為希望可以讓模型在學習過程提高績效，此外，我們也嘗試修改神經網路架構，像是修改 batch size、loss function 等調整超參數的方式，來降低模型訓練時出現太大的 gap generalization。

## 2.2 操作說明

本研究執行環境採用 Python3.9.13，以 Visual Studio Code 作為開發工具，利用 Tensorflow 及 Keras 和網路上公開的模型來建構神經網路，並使 Pandas、Numpy、PIL、cv2、Matplotlib 等函式庫來讀取圖片，資料前處理、以及模型預測出來的圖片和績效以視覺化的形式呈現，透過圖片壓縮、正規化和資料增強技術等增加模型泛化能力，在訓練模型時，用 ReduceLROnPlateau、EarlyStopping 等 callbacks 來動態調整超參數，最後使用邊緣檢測和形態學(侵蝕、膨脹)來更精準的計算我們自定義指標準確率。

## 三、實驗

### 3.1 資料集

名稱: Endotracheal Tube 資料集

Fold1 訓練資料筆數: 287

Fold1 測試資料筆數: 47

Fold1 驗證資料筆數: 47



圖 1 Fold1 氣管 X-ray 影像圖與 Mask 圖

Fold2 訓練資料筆數: 287

Fold2 測試資料筆數: 47

Fold2 驗證資料筆數: 47

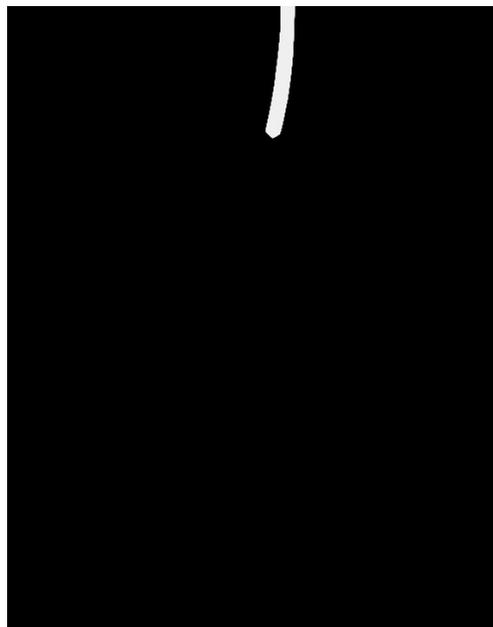
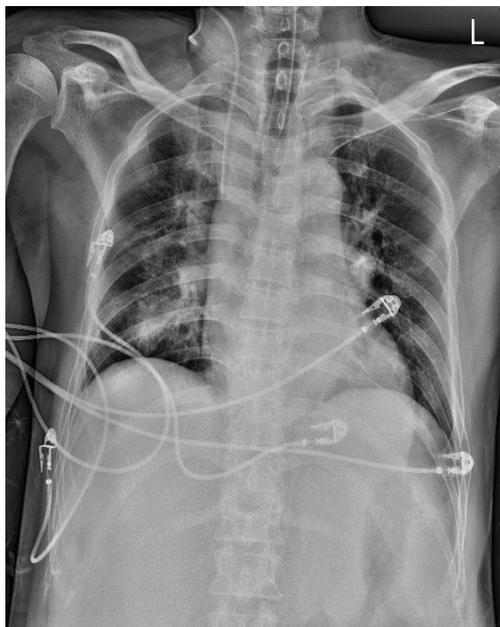


圖 2 Fold2 氣管 X-ray 影像圖與 Mask 圖

Fold3 訓練資料筆數: 287

Fold3 測試資料筆數: 47

Fold3 驗證資料筆數: 47



圖 3 Fold3 氣管 X-ray 影像圖與 Mask 圖

Fold4 訓練資料筆數: 285

Fold4 測試資料筆數: 48

Fold4 驗證資料筆數: 48

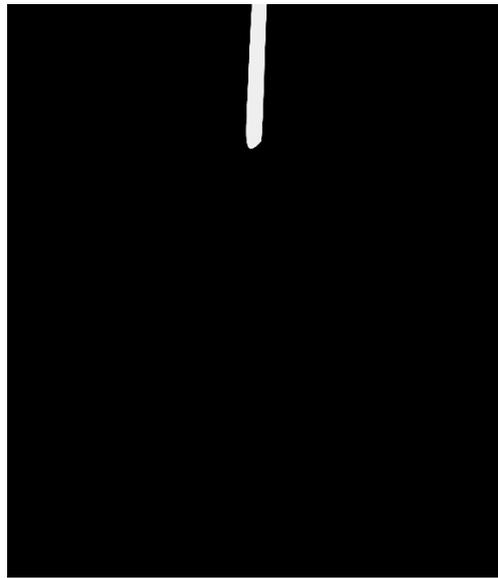


圖 4 Fold4 氣管 X-ray 影像圖與 Mask 圖

Fold5 訓練資料筆數: 285

Fold5 測試資料筆數: 48

Fold5 驗證資料筆數: 48



圖 5 Fold5 氣管 X-ray 影像圖與 Mask 圖

## 3.2 前處理

### ● 資料前處理

- 將所有氣管 X-ray 影像進行切割，原始影像的左右兩側以及下方分別裁切影像的  $\frac{1}{4}$ ，結果如下圖 6 右方影像所示。

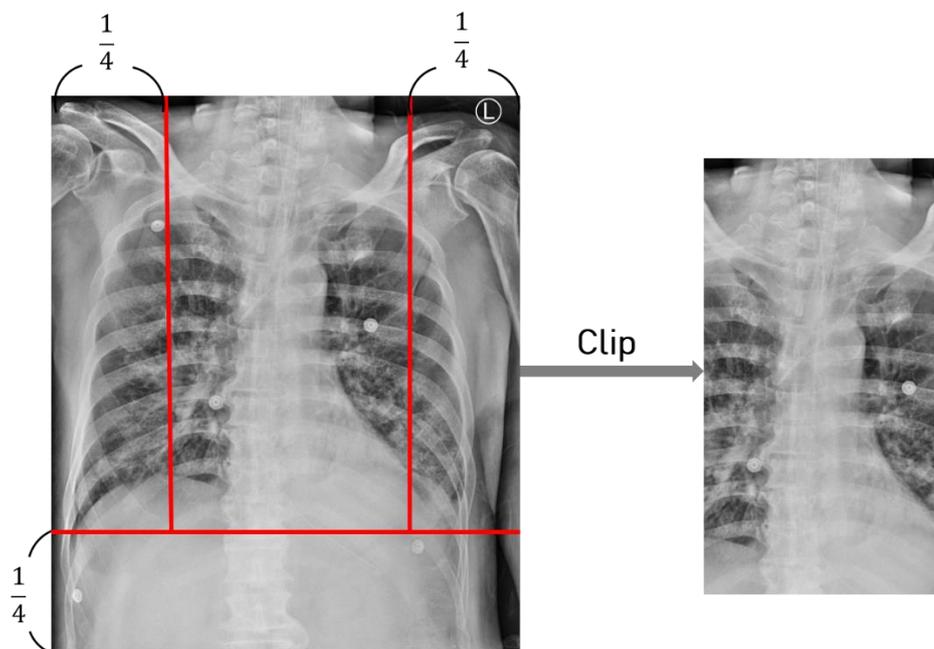


圖 6 氣管 X-ray 影像切割示意圖

- 將裁切後的影像進行 resize 成影像大小為  $320 \times 320$ 。
- 使用正則化技術(Regularization)針對所有氣管 X-ray 影像進行資料增強，以下為資料增強的細節說明。
  1. 對氣管 X-ray 影像隨機向上或向下以及向左或向右位移。
  2. 對氣管 X-ray 影像隨機進行水平旋轉。
- 將氣管的 mask 做二值化處理

## 3.3 實驗設計



圖 7 實驗設計流程圖

本研究實驗設計如上圖 7，把原始資料做裁減和壓縮得到一個比較小的 size，另外將 mask 做二值化處理，好讓模型比較好學習。在丟進模型之前我們也做了資料增強的動作，包括旋轉、上下左右位移和水平翻轉，希望模型可以有更好的泛化能力。我們模型選擇 TransUNet 和 AttentionUnet 做為我們模型架構並且評估績效預測氣管的遮罩，最後再算公分誤差率和 1.0cm、0.5cm 以內的準確率，我們用了邊緣檢測和形態學的方式解決圖片有可能有雜訊的問題。

## 3.4 實驗結果

### 3.4.1 氣管內管資料集(TransUNet)

首先我們使用了 TransUNet 架構，它結合了 Transformer 的自注意力機制，使模型能夠有效地捕捉全局上下文信息，同時利用 U-Net 的編碼器-解碼器結構實現了多尺度的特徵提取和重建。這種結合可以提高模型對細節和全局結構的理解能力，從而改善影像分割的結果，圖 8 為模型架構圖。

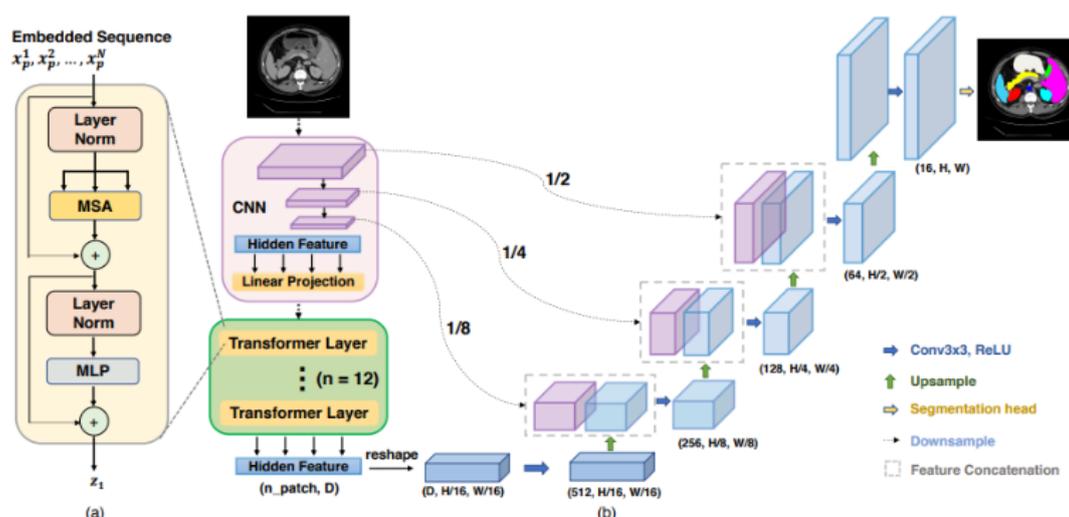


圖 8 TransUNet 架構

### Fold1 資料集

在 Fold1 資料集模型的超參數設定上面，我們除了調整模型裡面的超參數設定，也包括了 epoch、學習率的動態調整和損失函數選擇等超參數，其中模型裡面的激活函數是不動的包含 transformer 層是使用 Relu 激活函數，mlp 層是使用 GELU，最後一層輸出層是使用 Sigmoid，優化器都是使用 adam，mlp 的節點也固定為 512，另外 ReduceLROnPlateau 設定的學習率調整 patient 也固定為 3，我們總共測試 4 種超參數組合，如表 1 所示。

表 1 超參數設定的組合

參數設定 組合	Batch size	Early stopping (patient)	layers	embed_dim	num_heads	Reduce LR On Plateau (factor)	Loss function
Set_1	8	10	6	387	8	0.5	BCE_DICE_loss
Set_2	4	20	10	768	2	0.5	BCE_DICE_loss
Set_3	4	20	10	768	2	0.8	BCE_DICE_loss
Set_4	4	20	10	768	2	0.8	IoU

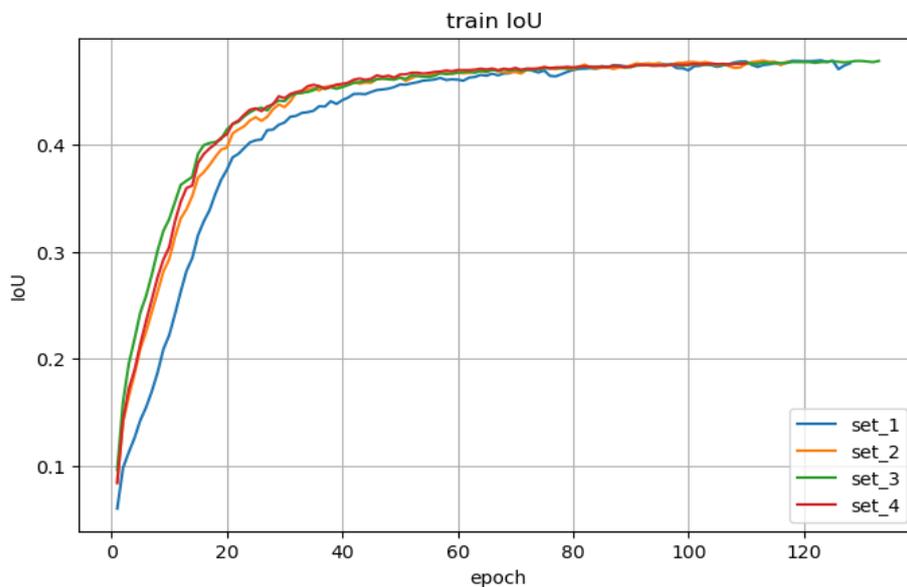
其中 IoU 計算的 loss 方式為：1-IoU

下表 2 為上述 4 種組合個別儲存最好的權重用來評估 4 種或 5 種指標績效，只有 IoU 是越高越好，其餘的指標都要是越低越好。

表 2 4 種組合個別的績效(驗證集為準)

	Loss(BCE_DICE_loss)	BCE	DICE	IoU	IoU_loss
Set_1	0.0474	0.0002	0.0597	0.4702	NaN
Set_2	0.0629	0.0002	0.0623	0.4689	NaN
Set_3	0.0472	0.0002	0.0599	0.47	NaN
Set_4	0.0553	0.0002	0.05862	0.4707	0.1102

由於我們主要著重於 IoU 指標，因為它是計算真實圖像和預測出來圖像的重疊程度，所以可以比較好了解模型的好壞，因此在 Fold1 資料集我們只顯示每個 epoch 對應的 IoU 值，如圖 9 所示為每個 Epoch 對 IoU 的對應關係。



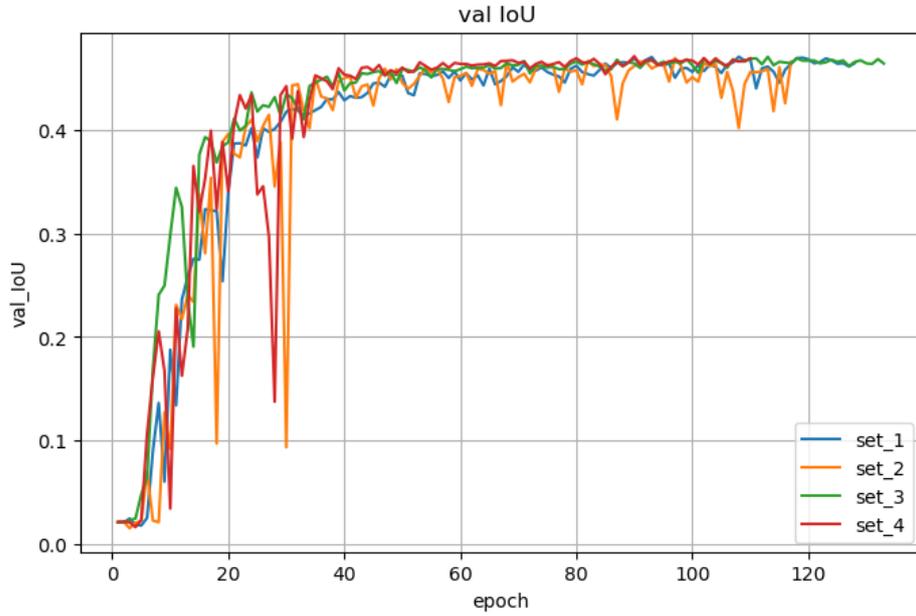


圖 9 IoU 不同超參數組合的績效

經過不同組合測試之後，我們選擇了 Set\_4 權重檔案做為我們後續的測試，圖 10 為 Set\_4 的 lr 和 IoU 的對應關係，下表 3 為 fold1 測試資料集，總共 47 筆的平均績效。

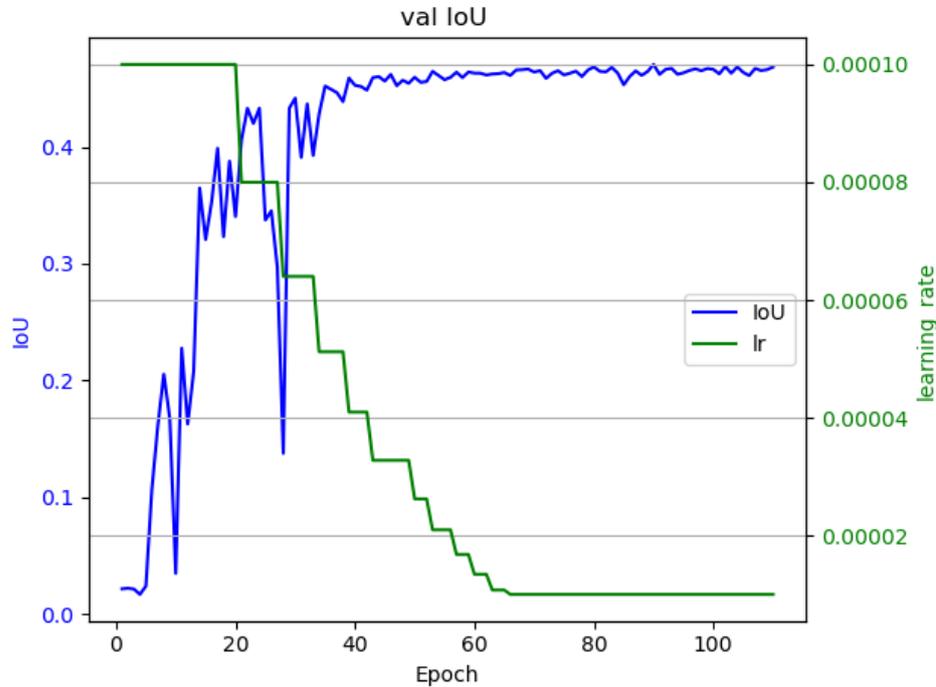


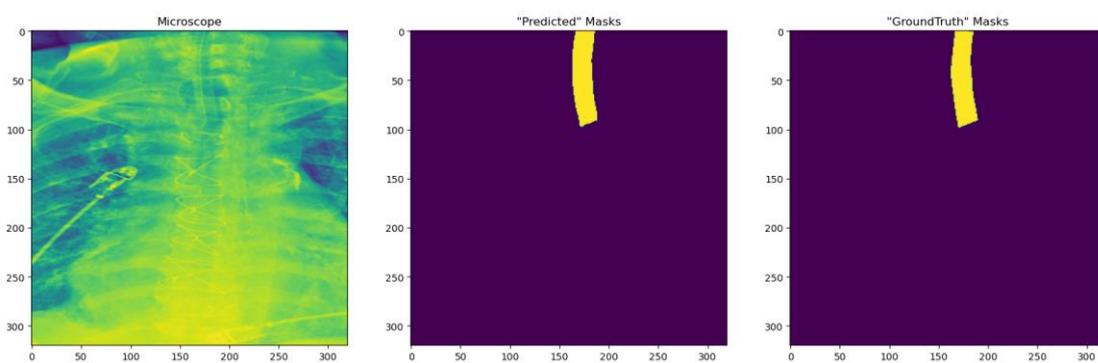
圖 10 IoU 和 lr 對應關係在每個 Epoch

表 3 set\_4 在測試資料集上的評估指標

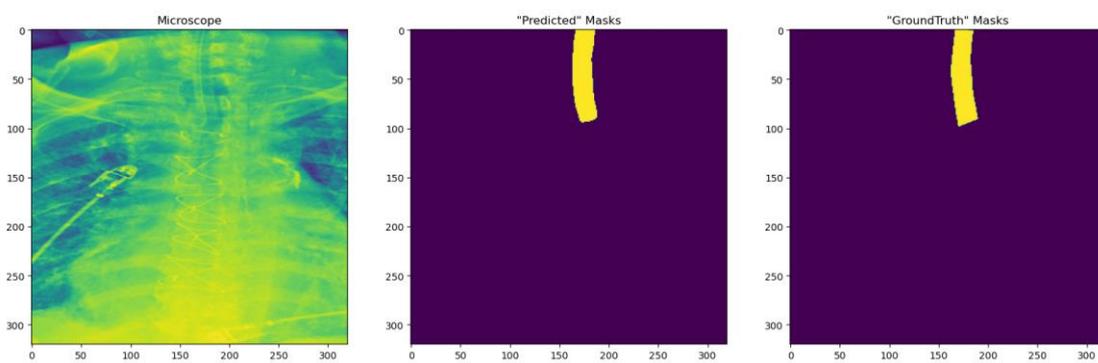
Loss(IoU_loss)	BCE DICE loss	BCE	DICE	IoU
0.1260	0.0677	0.0002	0.0676	0.4622

圖 11 顯示 Set\_1、Set\_2、Set\_3、Set\_4 的模型權重，分別來預測 fold1 的 test

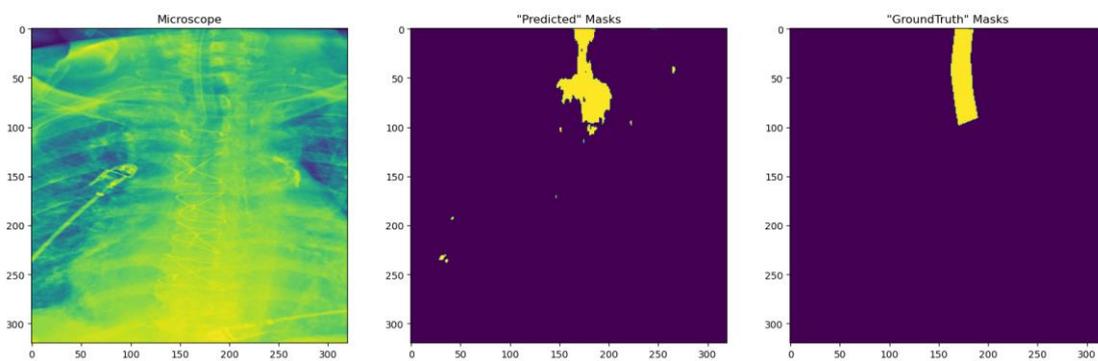
資料集裡第一張氣管內管的 mask。



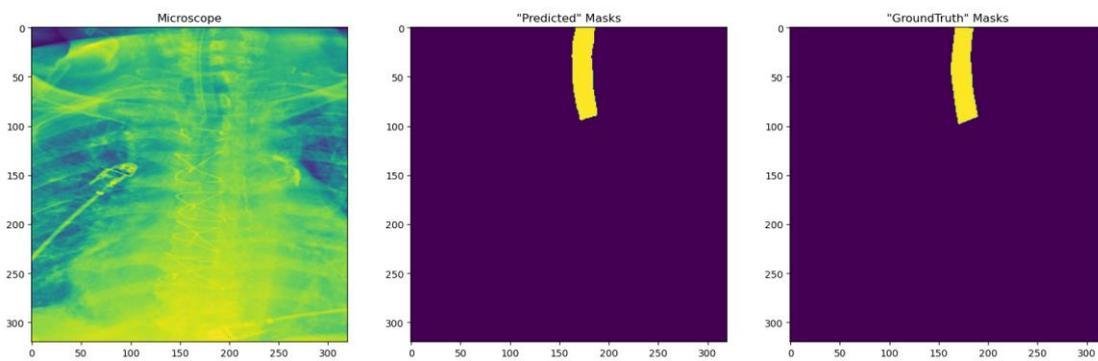
Set\_1



Set\_2



Set\_3



Set\_4

圖 11 fold1 測試集第一張的 4 種組合參數預測的結果

## 自定義指標計算

我們除了使用上面最常見的使用評估指標外，我們也自己建立了 3 種評估指標來衡量模型的好壞，分別計算 ETT 端點的預測位置包括：平均誤差公分、誤差在 0.5 公分內準確率、誤差在 1.0 公分內準確率，我們計算端點的方式是使用以 Y 軸為主，遍歷圖像的每一列直到找到有一列是沒有 mask 的(即那一列的像素值都為 0，就是都是圖像都是黑色)就可以大概算出氣管的長度。

由於上面算出來的長度是以 pixel 為基準，因此需轉成公分，我們以每 72 個 pixel 為 1 公分，下去計算測試資料的 mask 和我們預測出來的 mask 他們的長度差距有多少，若小於 72 代表它的誤差是在 1.0 公分以內的，那我們就視為是正確的反之就是錯誤的，來計算總體準確率有多少，0.5 公分也是如此。

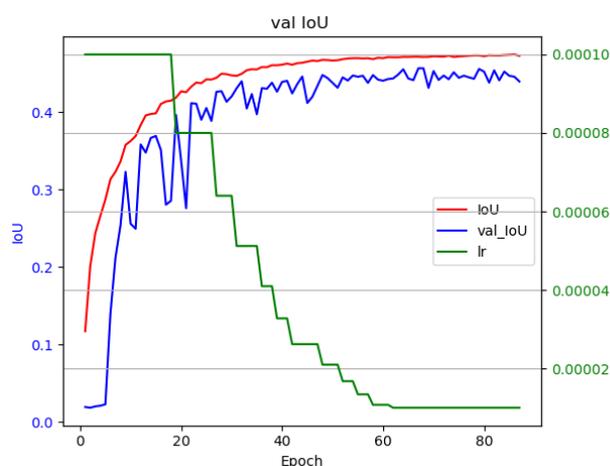
那我們計算誤差方面也做了一些邊緣檢測和型態學中的膨脹和侵蝕，來讓我們在計算誤差公分數的時候可以更精確的計算，不受其它雜訊所影響。

我們最終在 fold1 測試資料集上測出來的結果為：

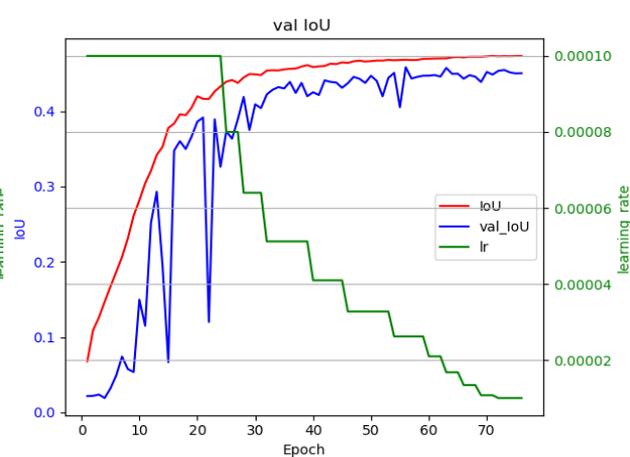
- 平均誤差公分：0.31cm
- 誤差在 1.0cm 內準確率：93.62%
- 誤差在 0.5cm 內準確率：80.85%

## Fold2-Fold5 資料集

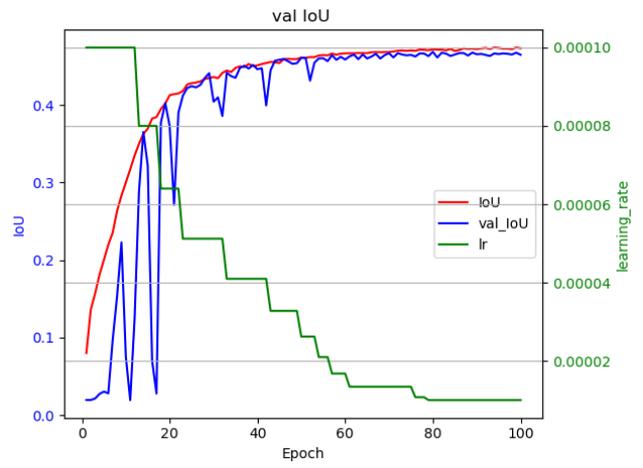
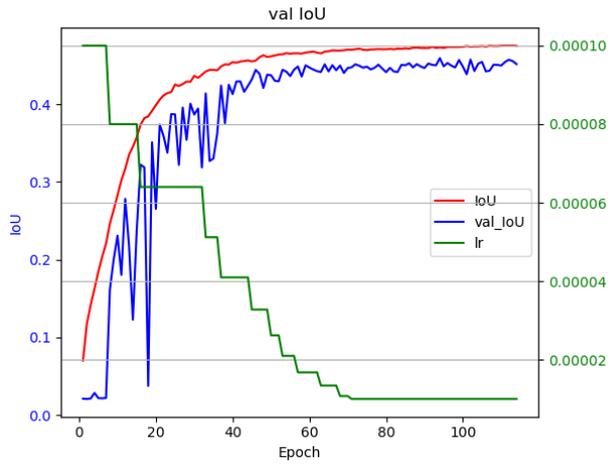
我們拿了在 fold1 中表現比較好的 set\_4 參數做為我們後續訓練 fold2 到 fold5 的模型參數，然後拿最好的權重檔用來測試個別資料集，圖 12 分別為不同的參數在驗證集上的訓練過程，以 IoU 為評估指標，圖 13 則是他們分別從各個資料夾裡面的第一張圖片來預測 mask。下表 4 為 fold2 到 fold5 在測試資料集上個別指標績效。



Fold2



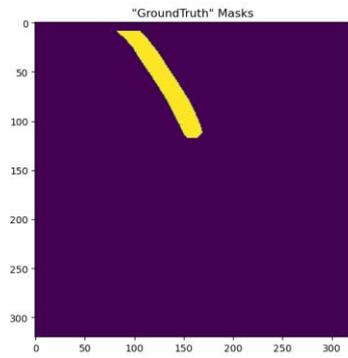
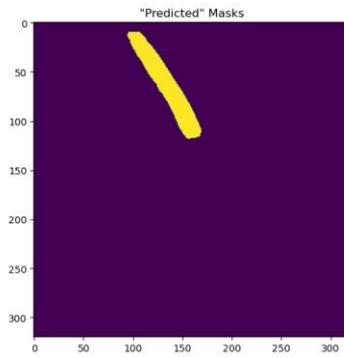
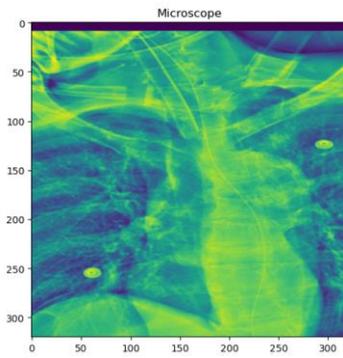
Fold3



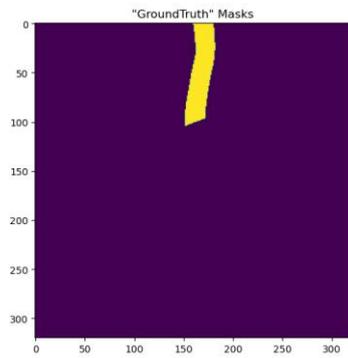
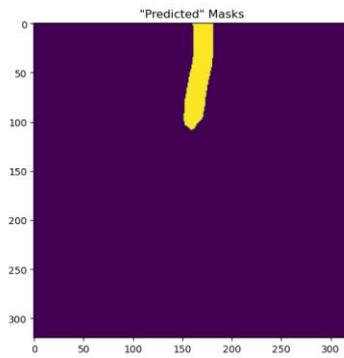
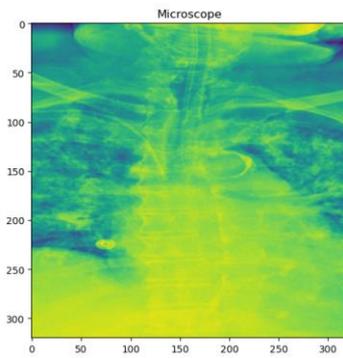
Fold4

Fold5

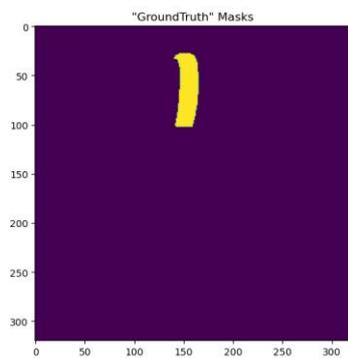
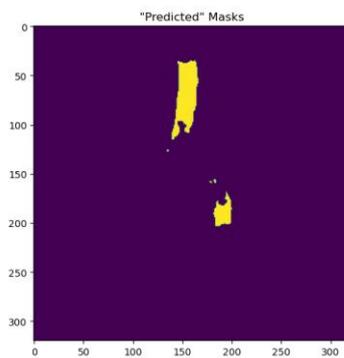
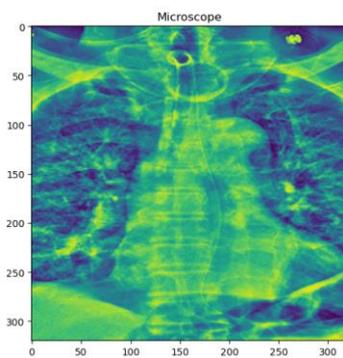
圖 12 Fold2-Fold5 IoU 和 lr 對應關係在每個 Epoch



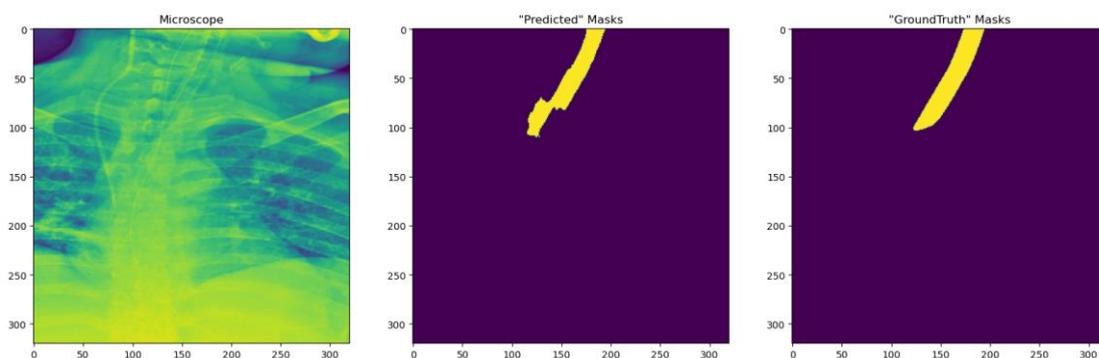
Fold2



Fold3



## Fold4



## Fold5

圖 13 Fold2- Fold5 測試集第一張的預測的結果

表 4 fold2 到 fold5 個別評估指標(測試集)

	Loss(IoU_loss)	BCE_DICE_loss	BCE	DICE	IoU
Fold2	0.1325	0.0718	0.0002	0.0713	0.4644
Fold3	0.2246	0.1358	0.0003	0.1361	0.432
Fold4	0.1739	0.0965	0.0003	0.0962	0.4519
Fold5	0.1717	0.0975	0.0003	0.0972	0.4514

我們也使用自己自定義的指標來衡量模型的績效，以下表 5 分別為 fold2-fold5 個別的平均誤差公分、誤差在 0.5cm 內準確率、誤差在 1.0cm 內準確率。

表 5 fold2 到 fold5 個別自定義指標(測試集)

	平均誤差公分	誤差在 1.0cm	誤差在 0.5cm
Fold2	0.40cm	93.62%	80.85%
Fold3	1.03cm	68.09%	46.81%
Fold4	0.54cm	87.5%	46.81%
Fold5	0.65cm	87.5%	77.08%

最後我們將所有資料集整合一起算總體評估績效，如下表 6。

表 6 fold1 到 fold5 總體平均績效(測試集)

總體 平均 績效	Loss (IoU_loss)	BCE_ DICE_ loss	BCE	DICE	IoU	平均 誤差 公分	誤差在 1.0cm (%)	誤差在 0.5cm (%)
Fold1 - Fold5	0.1657	0.0939	0.0003	0.0937	0.4524	0.684	82.184%	66.48%

## 結論

我們使用了一種用於圖像分割的深度學習架構，來預測氣管插入人體的那個位置，這個實驗我們使用了 transunet 的架構，他是結合了 unet 和 transformer 兩種模型下去組合的，我們最後所用的 set\_4 參數組合參數量高達 65,991,809，那訓練的結果我們可以看到 Fold1-Fold5 其實績效都是差不多的，只有 Fold3 相較起來差一點，我們也嘗試調了不同種的參數組合，但其實都影響不大，最後我們選擇 set\_4 作為後續測試，因為它 IoU 指標相對最好 0.4707，那它主要的改變是損失函數的計算，另外我們發現動態的學習率調整也可能讓模型有機會學得更好，從上面實驗結果可以發現有些模型適當的調整學習率，是可以讓驗證集突破現有的績效的。

雖然說在上述的各個指標中表現的並不是很好，但是如果我們實際拿出來看它在測試資料預測的圖片和我們自定義的指標來觀察的話，其實它表現的是不錯的，基本上只有幾張預測的圖片是不準的，其它的預測圖片都蠻像的，因此 IoU 會比預期的低的一種可能性是它圖像外觀是對的，但是在位置上可能並沒有學得很好，導致再算重疊的部分時重疊地方很少，這可能是我們需要再改進的地方，在自定義指標方面，尤於我們有做一些前處理像是邊緣檢測之類的方法等，因此得出的結果 Fold1、Fold2 和 Fold5 準確率其實都蠻高的，在 Fold3 和 Fold4 比較沒有這麼好，尤其誤差在 1.0 和 0.5 公分的準確率差蠻多的，但是總體來說除了 Fold3 的平均誤差超過 1 公分之外，其餘資料集都是誤差在一公分以內的。

## 參考文獻

Timur Abdualimov (2023).UNet++: Implementation of the UNet++ architecture on TensorFlow for segmentation of cell nuclei.

<https://medium.com/mllearning-ai/unet-implementation-of-the-unet-architecture-on-tensorflow-for-segmentation-of-cell-nuclei-528b5b6e6ffd>

yingkaisha (2022).keras-unet-collection.

<https://github.com/yingkaisha/keras-unet-collection>

SerendipityQYK (2022).TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation.

[https://blog.csdn.net/weixin\\_49627776/article/details/115710379](https://blog.csdn.net/weixin_49627776/article/details/115710379)