

一、 Introduction

腮腺是人體三對唾液腺中最大的一對，但腮腺腫瘤罕見，僅佔人類腫瘤的 0.3% 和頭頸部腫瘤的 6%。由於腫瘤的罕見性，文獻中並沒有太多報告。因此，我設計了一個系統，使用頭頸部 CT 影像數據來偵測腮腺和分割腫瘤類型。

這樣的系統對於醫學影像診斷領域可能具有重要意義。通過收集和分析頭頸部 CT 影像數據，可以幫助醫生更準確地識別腮腺腫瘤的類型和位置。這有助於制定更有效的治療計劃和提高患者的治療成功率。雖然目前相關文獻資料較為有限，但構建一個具備自動辨識和分割腮腺腫瘤功能的系統，將豐富這一領域的知識體系，為醫學研究和臨床實踐提供寶貴的支持。

準確檢測腫瘤非常重要，此系統可以分類出三種類型：Cancer、Mix 和 Warthin 腫瘤。目標是在操作過程中自動檢測並分割腫瘤。第一步從提供的圖像中偵測腮腺 (Left normal / Right normal)。第二步分割出有腫瘤的區域並且分類。

二、 Methods

1. 前處理

| | Cancer | Mix | Warthin | Total |
|------|--------|-------|---------|-------|
| 所有影像 | 240 張 | 240 張 | 240 張 | 720 張 |
| 測試集 | 3 張 | 3 張 | 3 張 | 9 張 |

- 將所有影像以 8:2 的方式，分成訓練集和驗證集。
- 將影像轉成 Tensor 型態。
- 將訓練集和驗證集切分為多個 batch，每個 batch 為一個訓練單位。

d. K-fold 驗證。

2. 模型

| Object detection | Segmentation | Classification |
|------------------|--------------|----------------|
| Yolov7 | U-net | ResNet50 |

3. 訓練參數

a. Object detection - Yolov7

- Epoch = 300
- Batch-size = 32
- loss = Compute Loss OTA
- Learning rate = 1e-2
- Optimizer = Adam
- Image size = 640 x 640

b. Segmentation - U-net

- Epoch = 100
- Batch-size = 8
- Loss = BCE With Logits Loss
- Learning rate = 1e-6
- Optimizer = RMSprop
- Image size = 224 x 224

c. Classification - ResNet50

- Epoch = 1000
- Batch-size = 16
- Loss = Cross Entropy Loss
- Learning rate = 1e-3
- Optimizer = Adam
- Image size = 64 x 64

4. 驗證

將所有影像以 8:2 的方式，分成訓練集和驗證集，訓練集總共 576 張影像，驗證集總共 144 張影像。

5. 計算指標

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

■ $Precision = \frac{TP}{TP+FP}$

■ $Recall = \frac{TP}{TP+FN}$

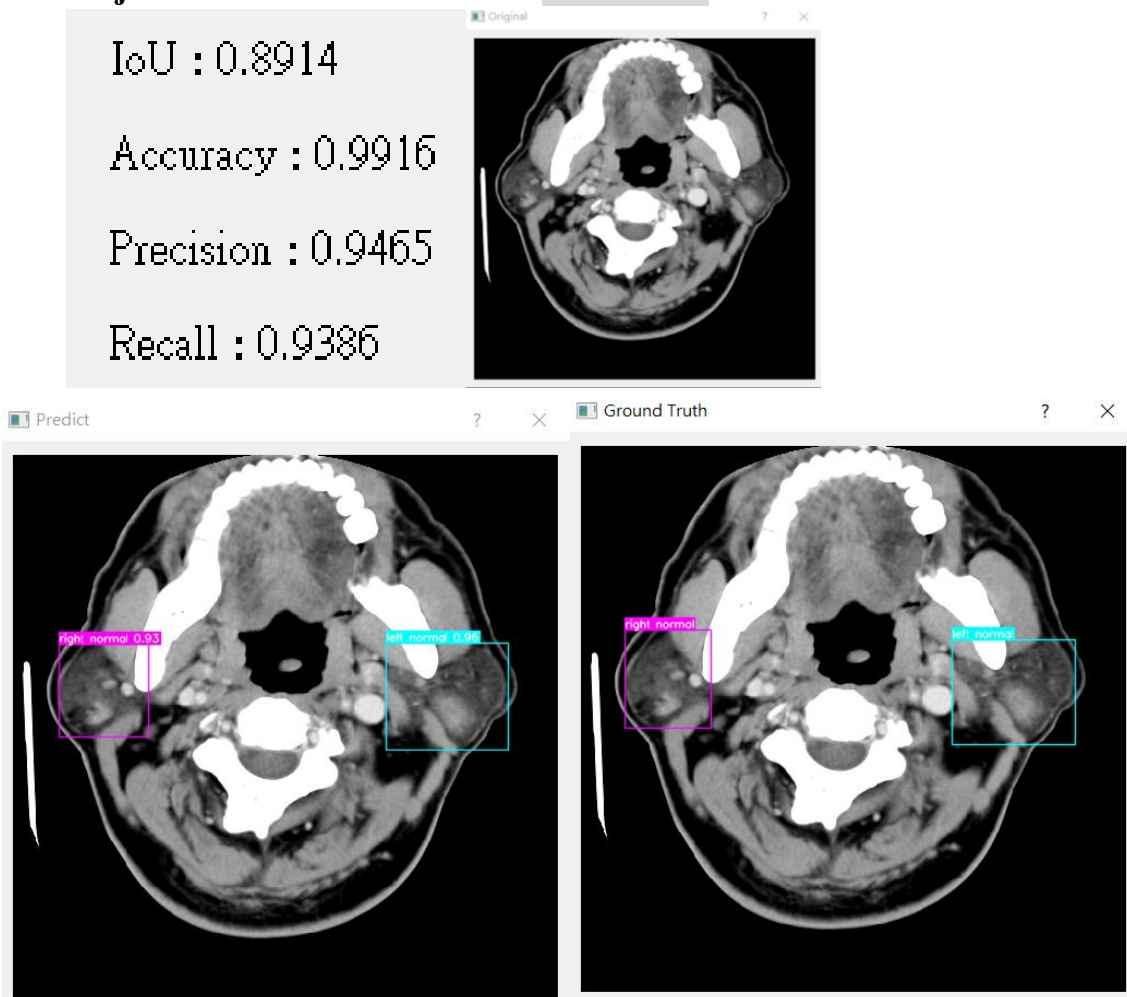
■ $IoU = \frac{A \cap B}{A \cup B}$

■ $Dice\ Coefficient = \frac{2 * |prediction| \cap |GT|}{|prediction| + |GT|}$

三、 Results

1. 總共分成四個部分，分別為：Objection detection best case、Objection detection worst case、Segmentation best case、Segmentation worst case。
2. 分類的顏色和 Segmentation 的結果同時呈現：Warthin - 紅，Mix - 藍，Cancer - 綠。

■ **Objection detection best case: IoU = 0.89**



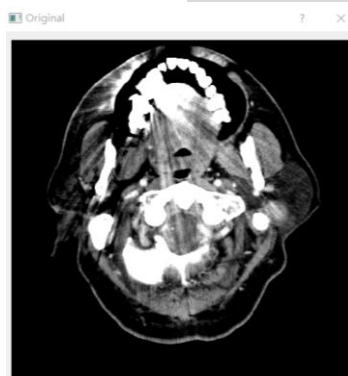
■ **Objection detection** worst case: $\text{IoU} = 0.40$

$\text{IoU} : 0.4021$

Accuracy : 0.9467

Precision : 0.4041

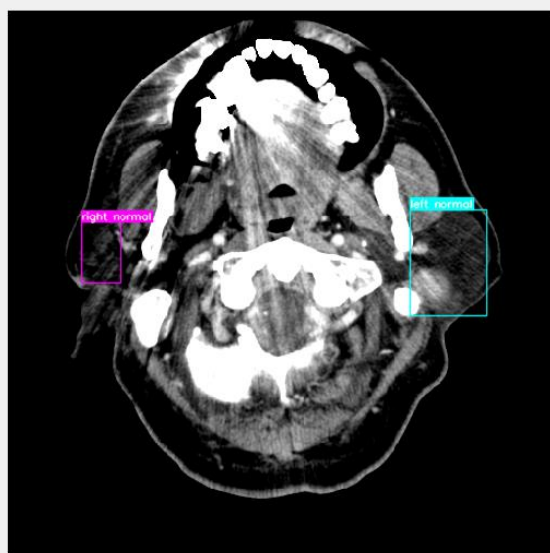
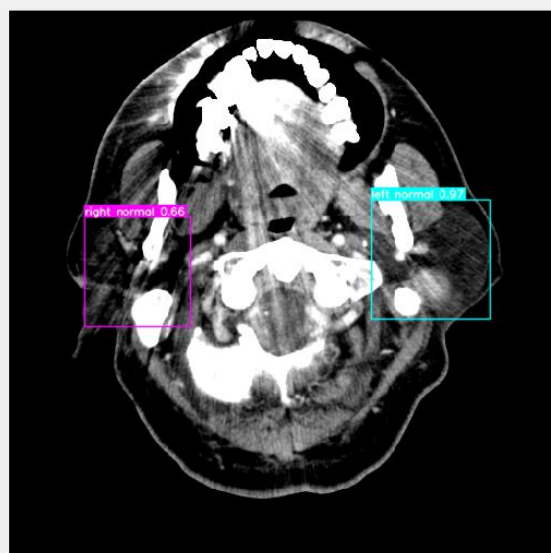
Recall : 0.9878



Predict



Ground Truth

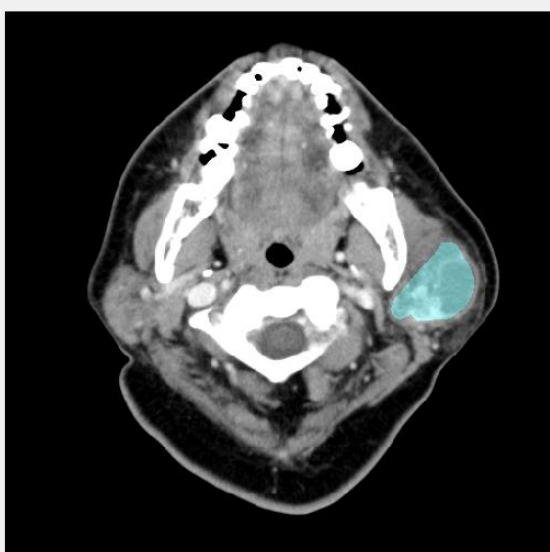
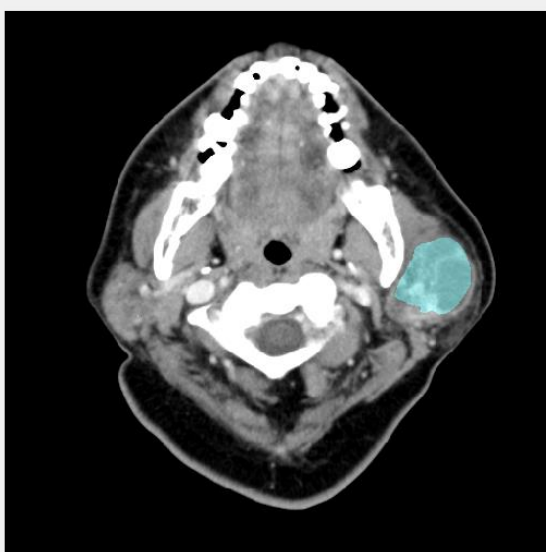


■ **Segmentation** best case: $\text{Dice} = 84.84\%$

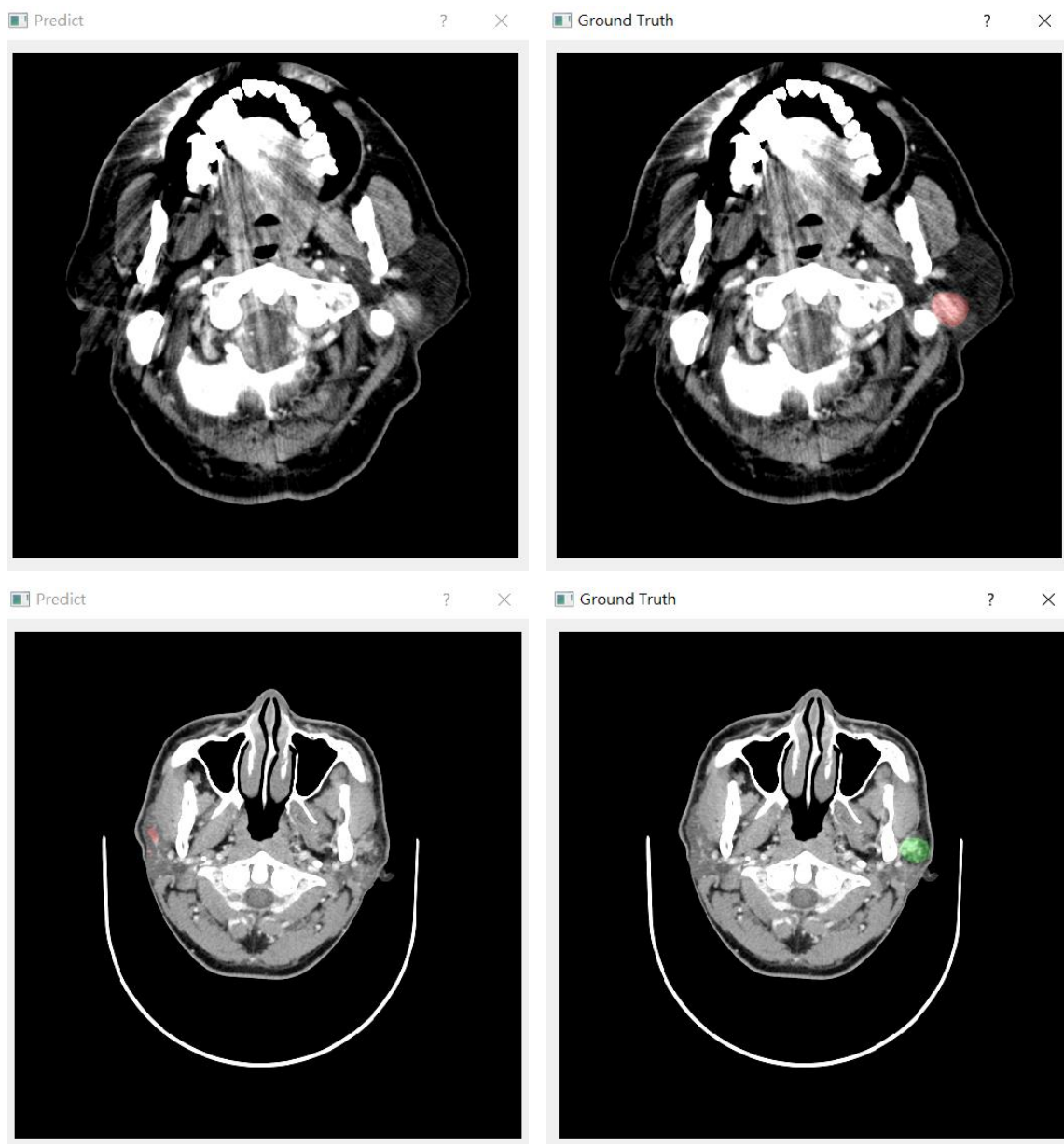
Predict



Ground Truth



■ Segmentation worst case: Dice = 0.00%



四、 Discussions

Object detection 最佳情況下的 IoU 為 0.89，而最差情況下為 0.40。這表示在最佳情況下，模型成功地準確識別了 Ground truth，並且檢測結果與 Ground truth 的重疊程度較高。然而，在最差情況下，IoU 較低，可能表示模型在某些情況下較難精確地定位目標，預測的腫瘤比 Ground truth 的腫瘤大，因此計算出來的 IoU 較低，但 Accuracy 仍然有 0.95。

Segmentation 的最佳情況下的 Dice 分數為 84.84%，而最差情況下為 0.00%。這表明在最佳情況下，Segmentation 模型能夠成功地將某些影像的 Ground truth 準

確地從圖像中分割出來，Dice 分數較高。然而，在最差情況下，Segmentation 失敗，可能由於影像本身不清晰，導致模型無法正確地捕捉目標邊界，此外，我發現 Segmentation 較差的影像也會發生分類錯誤的問題。

五、 Conclusion

這個系統結合了三個模型，首先使用 YOLOv7 來準確偵測左右腮腺，IoU 最高可達 0.89，接著使用 U-Net 進行精確的分割，Dice coefficient 最高達 0.85，最後利用 ResNet50 進行分類。這樣的整合方法讓系統能夠在製造過程中自動偵測並區分腮腺缺陷，並有效地將其歸類為 Cancer、Mix 或 Warthin 腫瘤。這種多層次的分析和識別有助於提高製造品質並及早發現潛在的問題，為後續的處理和治療提供了重要的指引。此系統的應用潛力和準確性為製造業和醫療領域帶來了極大的價值。