中國醫藥大學新竹附設醫院

醫學影像科

YOLOv7-based Automation for Parotid Tumor Detection

基於YOLOv7的腮腺腫瘤自動化偵測研究

王柏森

Bo-Sen Wang

指導教授: 阮春榮 博士

Advisor: ph.D.

中華民國 102年 12月

December 2023

**摘要**

論文名稱: 「基於YOLOv7的腮腺腫瘤自動化偵測研究」

單位:中國醫藥大學新竹附設醫院 醫學影像科

提要別:112年12月論文提要

員工:王柏森 指導教授:阮春榮 副院長博士

論文提要內容:

在本研究中，我們深入研究了YOLOv7在腮腺腫瘤CT影像上的應用，該方法將腮腺腫瘤CT影像作為輸入，利用YOLOv7模型實現腫瘤的自動化偵測。為了評估YOLOv7模型的性能，我們引入了四種機制，以確保提高對於腫瘤偵測的可信度和臨床解釋性。

實驗結果顯示，在我們提出的基於腫瘤為基準點的評估方式下，我們的模型能夠合理評估其在臨床應用中的性能，同時在數據層面呈現了一定的效果。這一評估方式為模型的臨床應用提供了實用且可靠的參考，為腮腺腫瘤的自動化偵測提供了更為準確且可靠的工具。我们的研究成果有望為未來臨床醫學影像分析和疾病診斷領域提供實用的指南。

關鍵字: computed tomography, parotid tumor, YOLOv7, Automation Object Detection.

**Abstract:**

Our study delves into the application of YOLOv7 for automated parotid tumor detection on CT images. Using CT scans as input, the YOLOv7 model is employed to achieve automated tumor detection. We introduced four metrics to evaluate the model's performance, ensuring increased reliability and clinical applicability in tumor detection.

Results indicate that our proposed evaluation method, centered around the tumor, allows the model to reasonably assess its clinical performance and demonstrates promising outcomes at the data level. This approach provides a practical and reliable reference for the clinical use of the model, offering a more accurate tool for parotid tumor detection. Our research contributes valuable insights to the field of clinical medical imaging and disease diagnosis.

Keyword: computed tomography, parotid tumor, YOLOv7, Automation Object Detection

**目錄**

**摘要**…………………………………………..…………….……**2**

**Abstract**………………………….……..…………………….…**3**

Outline………………………………..…………………….……**4**

**1. Introduciton…….……………………………….…………...5**

**2. Related work………….………………………………...……6**

2.1 Topics Related to Parotid Tumors………..……………6

2.2 Topics Related to Localization…………..…………….7

**3. Methodology……………………………….….……………..11**

3.1 Datasets and Data Processing…………..………..……11

3.2 The method………………………………..…………..12

3.3 Object Detection Evaluation Metrics……..…………...13

3.4 Evaluation………………………………………..…….14

**4. Experiment Analysis……………………………….………...16**

#### 4.1 Based on the Slices Space…………………………..….17

#### 4.2 Based on the Detection Box Space…………..…………18

4.3 Based on the Tumor Space…………………..………….19

4.4 experiment conclusion……………………..……………21

**5. Conclusion….…………………………………….…………….25**

**6. Reference………………………………………….……………25**

1. **Introduction**

自動化偵測腮腺腫瘤在醫學影像分析領域扮演著極其重要的角色。手動標記腮腺腫瘤CT影像的困難性不僅使過程耗時費力，更受制於標記者主觀意識和疲勞程度的影響。面對每位患者腮腺腫瘤的CT影像組，涵蓋大量切片使得標記變得極為耗時，同時也容易產生標記錯誤，降低了標記的可靠性。這樣的困難性不僅增加了臨床工作的負擔，更限制了腫瘤資料的應用範圍。考慮到一位醫生需觀測大量CT影像，容易出現漏判以及的誤判情況，自動化偵測提供了一個初步的結果，有效減輕了醫生人工肉眼判讀的負擔。

為解決手動標記的問題，我們提出了基於YOLOv7的自動化偵測方法，旨在有效標記腮腺腫瘤並提高標記的精確性。研究中，我們蒐集了腮腺腫瘤的CT影像，並進行了最大最小值正規化，以便更有效地進行模型訓練。為了全面評估我們方法的效能，引入了多種評估指標，包括Accuracy、Recall、Precision、Specificity和NPV。這樣的多指標評估方法不僅提供了對模型效能的全面了解，還有助於根據不同的臨床需求選擇最適合的評估標準。

總的來說，本研究的貢獻體現在提出了自動化偵測方法，並建立了一個多角度的評估框架，以更全面地應對複雜的腮腺腫瘤CT影像。這項研究不僅旨在提高標記的效率和準確性，同時也推動醫學影像分析的發展，為臨床診斷提供更可靠的輔助工具。

**2. Related work**

在這一章節中，我們將提供有關使用機器學習相關技術進行腮腺腫瘤檢測的當前最先進方法的概述。

2.1 Topics Related to Parotid Tumors

Emrah Gunduz等人[1] 提出了一種方法，旨在創建一種基於多參數的MRI深度學習(DL)的人工智慧方法，用於腮腺常見腫瘤的鑑別診斷。在他們的研究中採用了InceptionResNetV2 DL 模型和多數投票法對123位患者的MRI 影像進行腮腺腫瘤分類。他們提出的方法在非組合方法和非組合方法的準確率實現了86.43%和92.86%的令人印象深刻的準確性。

Dr. Merve ÖNDER等人[2] 開發一種使用 U-Net 架構在頭頸部 CT 影像上自動分割腮腺的演算法，並評估該模型的性能。結果發現，AI 模型在軸向 CT 切片中分割腮腺的 F1 分數、精確度和靈敏度均為 1。AUC 值為 0.96。這項研究表明，使用基於深度學習的人工智慧模型在軸向 CT 影像上自動分割腮腺是可能的。

在[3]中提出設計基於多模態磁振造影像(MRI)序列的深度學習模型，用於腮腺腫瘤自動分類，並改善臨床環境中的診斷決策。這種方法準確性為0.85，敏感度為0.9，特異性為0.84。

在[4]的一項研究中，在區分良惡性腮腺腫瘤時，當異常檢測（AD）與深度學習（DL）模型和L 2約束的softmax損失相結合時，可以透過少量不平衡數據獲得較高的診斷準確性 ，該研究在診斷性能方面，實現了（ROC-AUC = 0.86 和PR-ROC = 0.77）的量化數據。

[5]中的作者開發了改進的 ResNet 模型來對腮腺惡性腫瘤和良性腫瘤MR 影像進行分類。最終演算法在腮腺癌診斷和分期方面的準確率為82.18%，AUC為0.93。相同在[6]中的作者進行了一項研究，研究旨在評估深度學習模型在平掃電腦斷層掃描 (CT) 影像上區分惡性和良性腮腺腫瘤的診斷效率。研究中利用ResNet50在CT平掃影像上區分腮腺良惡性腫瘤，ResNet50對腮腺惡性腫瘤的診斷具有較高的準確性和敏感性；準確度、敏感度、特異度和 AUC 分別為 90.8%、91.3%、90.4% 和 0.96。同樣在Resnet上，作者[7]中也提出了基於 ResNet18 的 DeepLab v3+ 架構的準確率為0.96153。

2.2 Topics Related to Localization.

[Mohamed Nasor](https://ieeexplore.ieee.org/author/37394553200) 等人提出[8] 提出了一種有效的腦部 MRI 影像腫瘤分割方法。該技術基於多閾值和物件計數技術，可以透過消除其他類似結構來準確檢測腫瘤的位置。實現 三個計算性能參數的靈敏度、特異性和準確度值分別為0.9707、1.0000 和 0.9995。

# [Prabira Kumar Sethy](https://ieeexplore.ieee.org/author/37085487539) 等人[9]提出了使用影像處理技術進行腦腫瘤定位，此研究使用了閾值處理、區域過濾、邊界追蹤、區域合併和影像減法等影像處理工具來定位腫瘤在 3000 個 MRI 上的準確率為 93%。

[Burcu Selcuk](https://ieeexplore.ieee.org/author/37090053099) 等人[10]提出了一種使用 BR35h 資料集和 YOLOv8 演算法進行腦腫瘤檢測的綜合方法，研究的目標是採用最新版本的 YOLO（即 YOLOv8）創建一個模型，在給定的 MRI 影像中定位和檢測腦腫瘤。透過各種評估，建議的模型實現了平均精度（mAP）為 97.6% 的高效能。

Francesco Mercaldo 等人[11] 透過提出一種用於檢測和定位腦癌的自動化方法，介紹了一種旨在檢測和定位腦癌的方法，該方法旨在增強腦癌的檢測和精確定位，以改善患者的預後和治療結果。利用 YOLO 模型自動檢測和定位腦癌在腦癌檢測中獲得了 0.943 的精確度和 0.923 的召回率。

**2. related work**

In this section, we provide an overview of the current stage-of the art methods in the field of tumor detection using machine learning techniques.

2.1 Topics Related to Parotid Tumors

Emrah Gunduz et al. [1] proposed a method aimed at developing an artificial intelligence approach based on multi-parameter MRI deep learning (DL) for the discriminative diagnosis of common tumors in the salivary glands. In their study, the InceptionResNetV2 DL model and a majority voting method were employed to classify MRI images of 123 patients with salivary gland tumors. The proposed method achieved impressive accuracies of 86.43% and 92.86% for non-ensemble and ensemble approaches, respectively.

Dr. Merve ÖNDER et al. [2] developed an algorithm utilizing the U-Net architecture for automatic segmentation of the salivary glands in head and neck CT images and assessed the performance of the model. The results revealed that the AI model achieved an F1 score, precision, and sensitivity of 1 in segmenting the salivary glands on axial CT slices. The AUC value was 0.96. This study demonstrates the feasibility of automatically segmenting the salivary glands on axial CT images using a deep learning-based artificial intelligence model.

In [3], a deep learning model is proposed based on multi-modal magnetic resonance imaging (MRI) sequences for automatic classification of salivary gland tumors, aiming to enhance diagnostic decision-making in the clinical environment. The method achieves an accuracy of 0.85, with a sensitivity of 0.9 and specificity of 0.84, showcasing its potential for improving diagnostic precision in clinical practice.

In a study presented in [4], a combination of anomaly detection (AD) with a deep learning (DL) model and L2-constrained softmax loss was employed for distinguishing between benign and malignant salivary gland tumors. This approach, utilizing a small amount of imbalanced data, achieved higher diagnostic accuracy. The study reported quantitative data on diagnostic performance, with a ROC-AUC of 0.86 and PR-ROC of 0.77.

In [5], the authors developed an enhanced ResNet model for the classification of malignant and benign salivary gland tumors in MR images. The algorithm achieved an accuracy of 82.18% and an AUC of 0.93 in the diagnosis and staging of salivary gland cancer. Similarly, in [6], the authors conducted a study to assess the diagnostic efficiency of deep learning models in distinguishing between malignant and benign salivary gland tumors on non-contrast-enhanced computed tomography (CT) images. Utilizing ResNet50 for the differentiation of salivary gland tumors on CT scans, the model demonstrated high accuracy and sensitivity for diagnosing malignant tumors, with accuracy, sensitivity, specificity, and AUC values of 90.8%, 91.3%, 90.4%, and 0.96, respectively. Furthermore, in [7], the authors proposed a ResNet18-based DeepLab v3+ architecture with an accuracy of 0.96153.

2.1 Topics Related to Localization.

In [8], Mohamed Nasor et al. proposed an effective brain MRI tumor segmentation method. The technique, based on multiple thresholds and object counting, accurately detects the tumor's location by eliminating other similar structures. The achieved sensitivity, specificity, and accuracy values for three performance metrics are 0.9707, 1.0000, and 0.9995, respectively.

In [9], Prabira Kumar Sethy et al. proposed a brain tumor localization method using image processing techniques. The study achieved a tumor localization accuracy of 93% on 3000 MRI scans, employing image processing tools such as thresholding, region filtering, boundary tracking, region merging, and image subtraction.

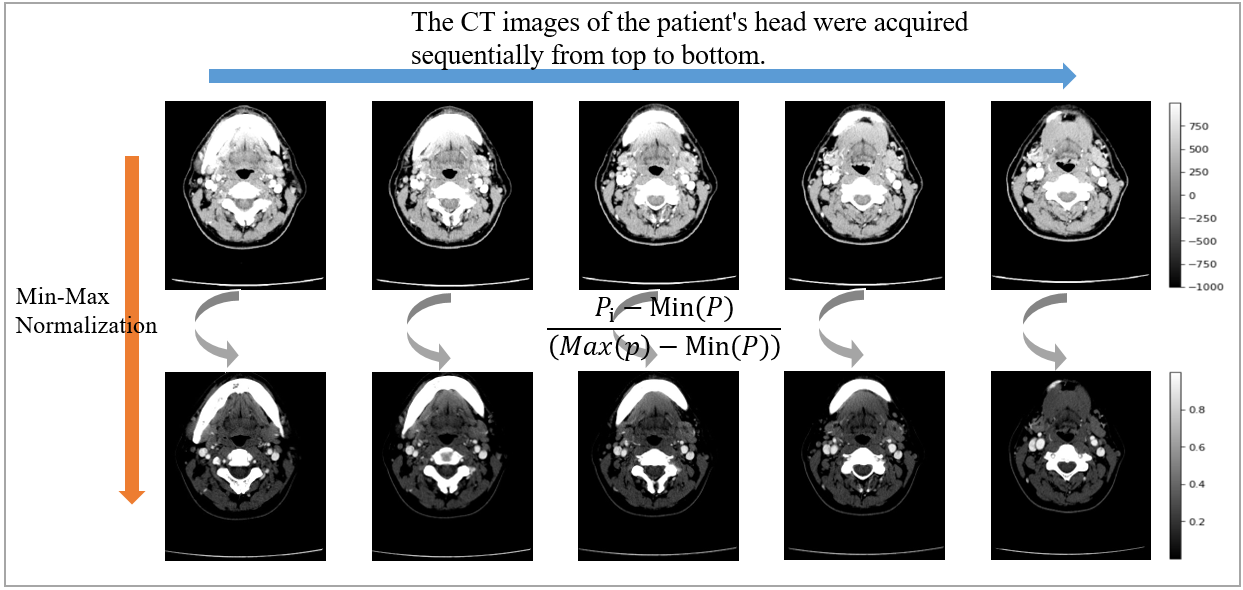
In [10], Burcu Selcuk et al. proposed a comprehensive approach for brain tumor detection using the BR35h dataset and the YOLOv8 algorithm. The research aimed to create a model with the latest version of YOLO (i.e., YOLOv8) to locate and detect brain tumors in given MRI images. Through various evaluations, the suggested model achieved high performance with an average precision (mAP) of 97.6%.

In [11], Francesco Mercaldo et al. introduced an automated method for the detection and localization of brain cancer, aiming to enhance the detection and precise localization of brain tumors to improve patient prognosis and treatment outcomes. Using the YOLO model, the automatic detection and localization of brain cancer achieved an accuracy of 0.943 and a recall of 0.923

**3. METHODOLOGY**

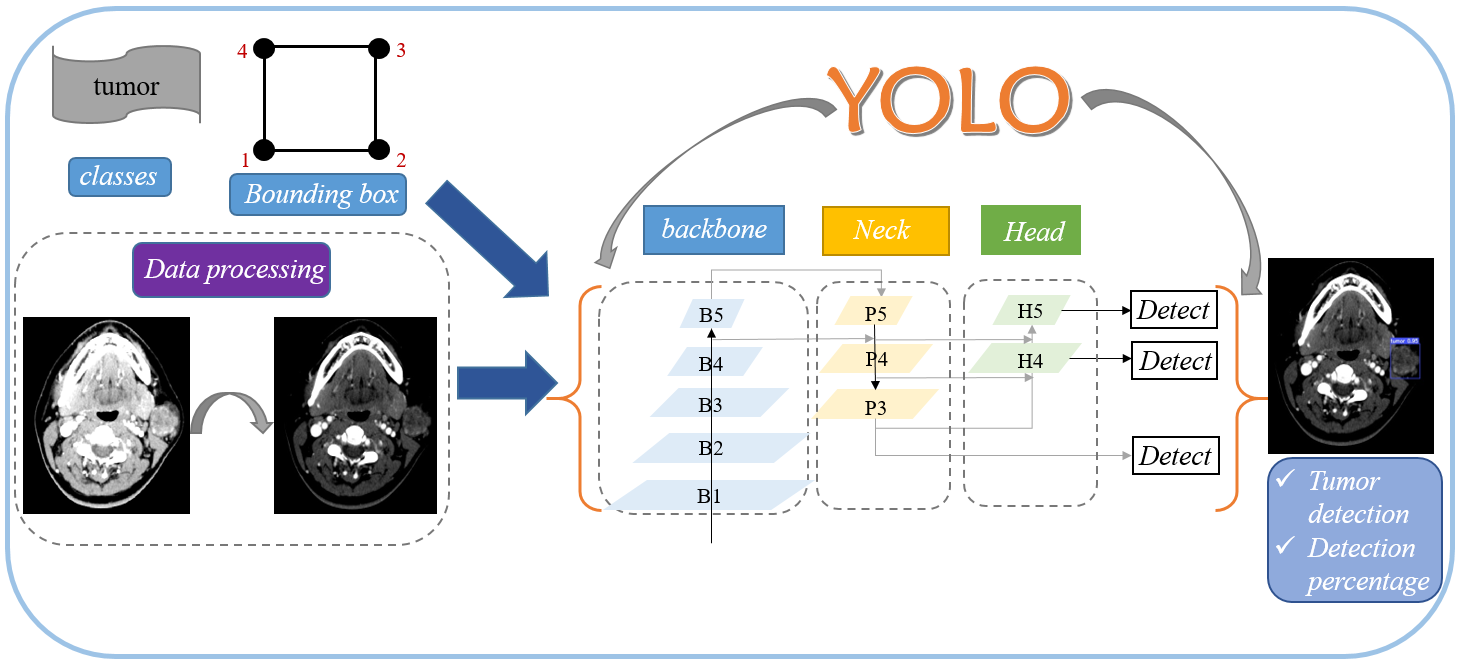
3.1 Datasets and Data Processing.

我們採用了中國醫藥大學新竹附設醫院的患者資料，透過電腦斷層掃描(CT)的方式，利用X光的光束穿透患者身體。在接收器收集穿透過身體的X光後，透過電腦解讀顯像，每位患者產生了一組近100張的成像，這100張CT影像由患者頭部到肩膀連續呈現，具有上下文關聯性。我們共擁有300組這樣的影像，其中45組作為訓練資料，其餘255組作為驗證集以評估模型效能。

鑒於YOLOv7無法處理整組影像作為3D輸入，我們將單一CT影像視為模型的輸入，並輸出模型判斷的腫瘤位置座標。由於CT影像的像素數值通常介於-1000至1000之間，與一般影像有所不同，這樣大範圍的數據可能導致模型難以收斂。因此，我們對單一CT影像進行了最大最小值正規化(Min-Max Normalization)，將其數值縮放至0~1之間(圖1)，以促使模型在訓練時更快達到收斂。

**Figure 1.** illustrates the visualization of raw data and normalized training data along with their corresponding ranges.

3.2 The method

在本章節中，我們描述了一種用於分析電腦斷層影像(CT)實現自動檢測和定位腮腺腫瘤的方法。本文的重點放在腮腺腫瘤的電腦斷層影像上，但所提出的方法也可以應用在其他類型的生物影像，類似的技術也可以應用在腦癌的腦部磁震造影(MRI)或其他器官的相關影像上，圖2.呈現用於腮腺腫瘤檢測以及定位方法的流程。

**Figure 2.** The overall structure of the proposed method for parotid tumor detection and localization.

為了建立一個有效的深度學習模型，用於從CT影像中掃描檢測和定位腮腺腫瘤，至關重要的是有一個包含腮腺腫瘤CT影像的相應標籤以及邊界框座標，意旨腮腺腫瘤定位的數據集。我們會強調擁有一個高品質資料集的重要性，其中包含CT影像和腫瘤定位的標籤。這樣的數據集為有效的訓練目標偵測模型奠定良好的基礎，透過準確的標籤，模型可以學習定位腮腺CT腫瘤存在的特定區域。

為了建構一個高效且準確的模型，我們將腮腺腫瘤影像先經過3.1章節的前置處理且將所有圖像調整為相同的尺寸，以確保模型能夠有效的收斂，在獲取了腮腺的CT圖像以及對應的標籤和邊界框訊息，下一步則是開發深度學習模型。

在摘要中提到，我們選擇採用YOLO模型[16-19]，這是一種先進的一階段深度學習偵測器，具備在影像中同時執行影像分類和準確物體偵測的能力。 YOLO的獨特之處在於能夠獨立完成整個過程（One Stage），輸入為一張影像，輸出包含兩個部分：一個邊界框向量和類別預測。 相較於現有的目標偵測模型，YOLO已被證明更為迅速。 這得歸功於YOLO在單一階段內執行辨識任務，與將特徵提取和物體分類分為獨立階段的其他模型不同。 YOLO直接預測輸入影像中存在的邊界框、物體機率和物體類別，這種簡化的方法使得YOLO在速度上具有顯著優勢，廣泛認為是目標偵測領域最出色的捲積神經網路模型之一。

值得注意的是，YOLO模型存在多個版本。 在本文中，我們選擇並實作了YOLOv7。

3.3 Object Detection Evaluation Metrics.

評估指標對於評估電腦視覺相關任務的性能是非常重要的，在目標偵測中有幾個被廣泛使用的評估指標，包含準確率(Accuracy)、敏感度(Sensitivity)又稱陽性預測值(TPV)、精確度(Precision)、特異度(Specificity)、陰性預測值(NPV)、F1分數(F1-Score), 交並比(IoU)。

準確率為所有樣本中，預測正確的比例，該指標顯示模型的綜合能力。

(1)

而敏感度則代表在所有陽性樣本中，正確預測為陽性的樣本比例，該指標在醫療領域中十分重要，它代表著漏檢測的風險這在醫療領域上是具有較大的代價，在統計學上我們又稱之為型一誤差(Type I error)。

(2)

精確度則衡量陽性預測的準確性，通過測量正確預測的陽性樣本佔所有陽性預測的比例來實現。該指標顯示模型減少偽陽性的能力。

(3)

特異度是指實際為陰性的樣本中，判斷為陰性的比例（例如真正未生病的人中，被醫院判斷為未生病者的比例）。該指標顯示模型減少偽陽性的能力。

(4)

陰性預測值是指在所有陰性樣本中，被正確預測為陰性樣本的比例。

(5)

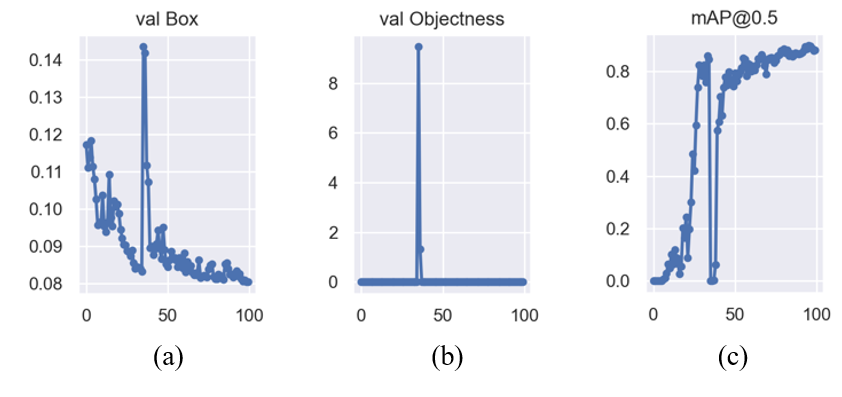
目前在辨識、偵測相關的[演算法](https://zh.wikipedia.org/wiki/%E7%AE%97%E6%B3%95)中經常會分別提到精確率（precision）和召回率（recall），F-score能同時考慮這兩個數值，平衡地反映這個演算法的[精確度](https://zh.wikipedia.org/wiki/%E7%B2%BE%E7%A2%BA%E5%BA%A6)。

(6)

目標偵測中使用的概念，是一種測量在特定資料集中檢測相應物體準確度的一個標準。 IOU表示了產生的候選框（candidate bounding box）與原始標記框（ground truth bound）的交疊率或說重疊度，也就是它們的交集與並集的比值。

(7)

3.4 Evaluation.

****在訓練過程中模型在每一次迭代都對驗證集進行全面評估，以確保有效的訓練，在這種情況下，驗證集包含來自數據集的27500張CT影像，訓練過程中，也同步觀測驗證集的偵測框回歸損失(bounding box loss)在最後一次迭代為0.08（圖3a）、目標檢測損失(objectness loss) 在最後一次迭代為0.0053（圖3b），並用於計算特定IoU閾值下的平均精度（mAP）。結果顯示，模型在訓練的最後一個時期在IoU = 0.45時實現了97.6％的mAP（圖3c）。腮腺腫瘤偵測模型針對臨床合理性進行全面評估，考慮以不同的角度了解模型的整體性能。

**Figure 3.** Results of the YOLOv7 on the validation datasets.

表明模型在每個時期中逐漸提高在檢測和定位腦腫瘤方面的性能（圖4a,4b）。這表明隨著訓練的進行，模型的性能變得越來越準確和可靠，值得一提的是在檢測損失大約第30次迭代時驗證損失有突然升高的現象，但隨著迭代次數增加也趨近於穩定。因此，對偵測框回歸損失、目標檢測損失的低損失值的結合表明，該模型在檢測和定位腦腫瘤以及在訓練過程中有效處理可變形特徵方面取得了顯著進展。

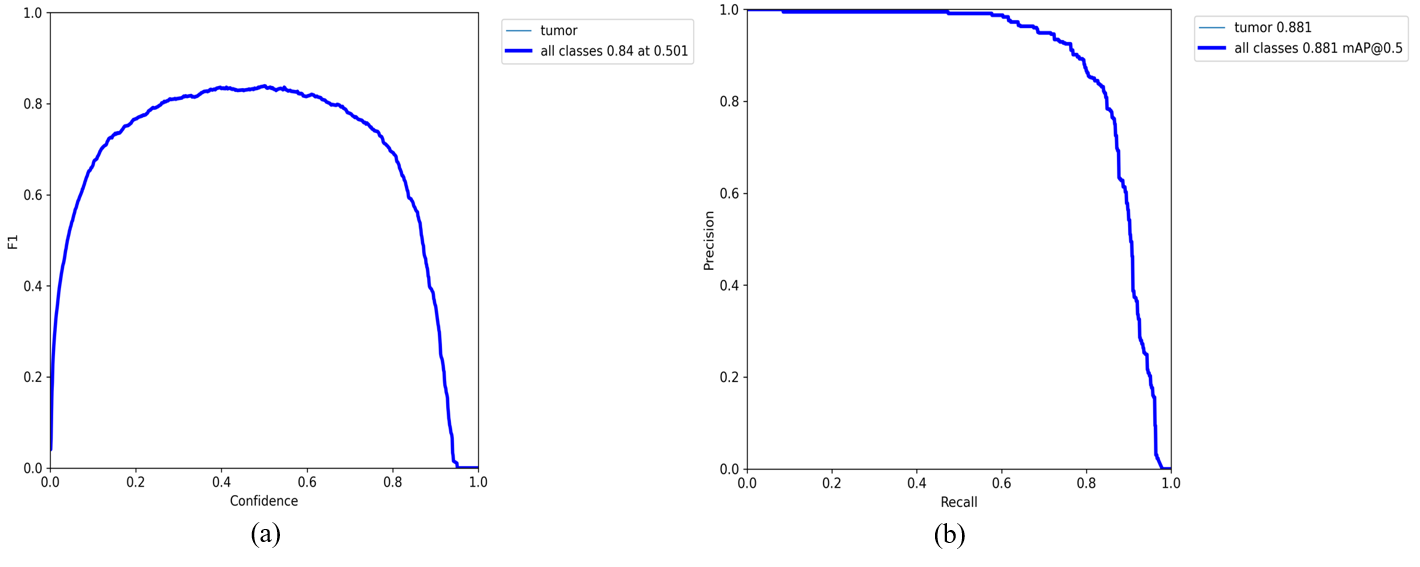
****精確度-召回曲線通常呈單調遞減趨勢，反映在目標檢測中精確度和召回率之間的權衡。提高精確度可能降低召回率，反之亦然。這種權衡來自於將更高閾值設為正預測的分類，提高精確度但降低召回率。相反，降低閾值以包含更多預測會提高召回率，但可能增加假陽性降低精確度。然而，精確度-召回率圖中可能存在異常，受類別不平衡、或數據集特性等因素影響。儘管典型趨勢是遞減，但應仔細分析每種情境的具體特徵。從圖4(b)中可見，此圖呈遞減趨勢，並顯示腦腫瘤檢測mAP@0.5相關的AUC為0.881。相同的F-score能同時考慮這兩個數值，平衡地反映這個演算法的[精確度](https://zh.wikipedia.org/wiki/%E7%B2%BE%E7%A2%BA%E5%BA%A6)給出更加適合的閾值(圖4a)。

Figure4 . Results of the YOLOv7 on the validation datasets.

在評估階段，我們將在第4章節對驗證集以及YOLOv7算法開發的腮腺腫瘤偵測模型針對臨床合理性進行全面的評估，考慮了以不同的角度了解模型的整體性能。

## **4. Experiment Aanlysis**

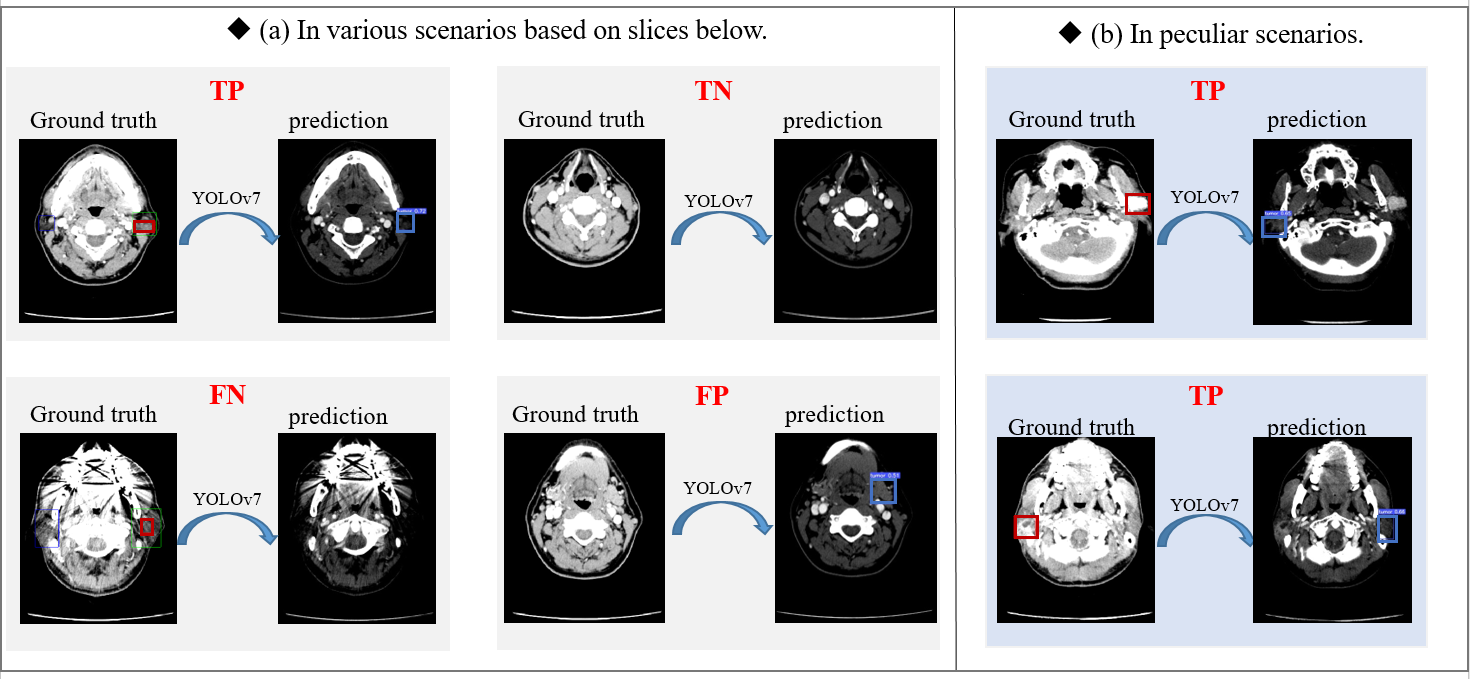
臨床上，病患頭部通過電腦斷層掃描(Computed Tomography，簡稱CT）掃描獲取斷層影像，並利用窗口技術生成相應組織的影像。這種三維影像的生成使得每位病患的頭部影像包含大約100張切片(Slice)。而傳統的腮腺腫瘤診斷方法是手動分析 CT 影像。然而，手動分析 CT 影像是一件費時費力的工作，而且容易出現誤診，因此我們採用YOLOv7自動化偵測(auto-detection)可以自動地從 CT 影像中檢測出腮腺腫瘤，從而提高診斷效率和準確性。而在處理這樣的數據時，我們面臨了以下評估腮腺腫瘤物件偵測模型效能的挑戰。

* 在進行腮腺腫瘤的物件偵測過程中，模型會生成一個偵測框(bounding box)，該框用於指示模型認為腫瘤位於影像中的位置，而我們會引入IoU，當作指標衡量偵測框是否正確檢測到腮腺腫瘤。
* 病患頭部通過電腦斷層掃描將會獲取一個立體三維影像，如何利用上下文的資訊制定一個合理且有效率的評估準則判斷病患是否罹患腮腺腫瘤，也是我們面臨的挑戰之一。
* 在腮腺腫瘤的偵測領域，我們需要一套設計良好的評估方式，以有效降低病患被錯誤標記為未罹患腮腺腫瘤的機會，這種誤判在統計學上被稱為型一誤差(Type I error)。在醫學上，型一誤差是一種具有相當高代價的誤判，因此設計出合適的演算法和評估準則以有效避免此類誤判，成為我們面臨的重要挑戰之一。

因此我們提出以下的三種評估方式評估模型效能。

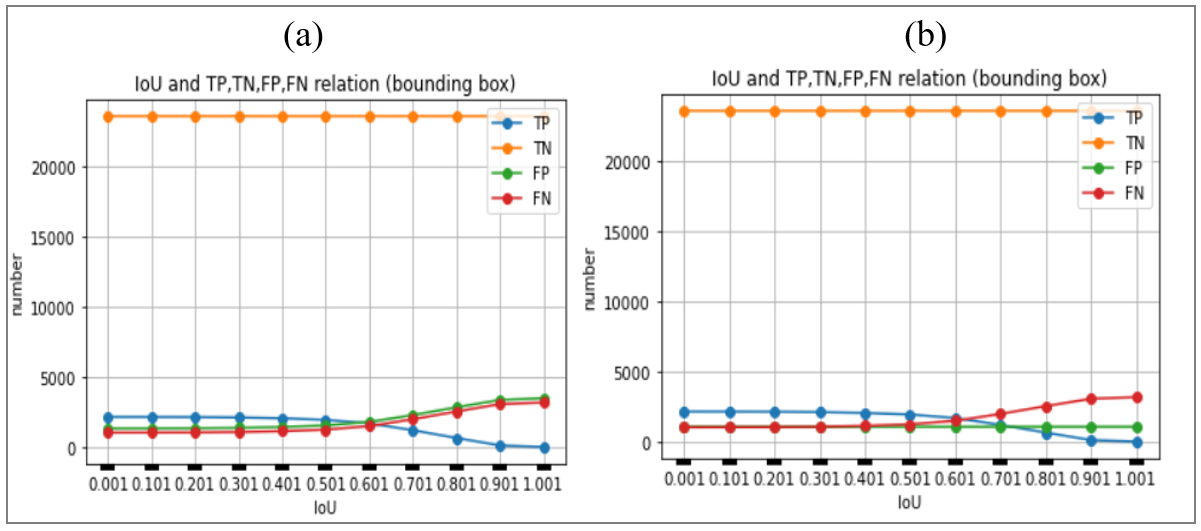
#### Based on the Slices Space(BS).

在切片空間的評估機制中，我們將每張影像視為一個切片(Slice)，對腮腺腫瘤的檢測進行判定。若單一影像中有腮腺腫瘤被檢測到，我們將其判定為真陽性(True Positive)；反之，若有腫瘤存在但未被檢測到，則視為偽陰性(False Negative)。對於單一影像中未存在腮腺腫瘤的情況，若模型未檢測到腫瘤，我們將其視為真陰性(True Negative)；相反地，如果模型錯誤地檢測到腫瘤，則視為偽陽性(False Positive) (圖2(a))。

然而，這樣的機制可能面臨一側腮腺腫瘤被檢測，但模型檢測框卻錯誤標記在相反側的情況。在這種情況下(圖2(b))， 根據目前的評估準則，將可能被錯誤地歸類為真陽性(True Positive)，導致不準確的評估結果。為解決此問題，我們需要進一步探討改進的評估機制，可能包括引入額外的上下文信息、考慮鄰近切片的情況，以及採用更複雜的演算法來提高腮腺腫瘤檢測的精確性。這樣的改進將有助於減少誤判的發生，提高評估機制的可信度和準確性。

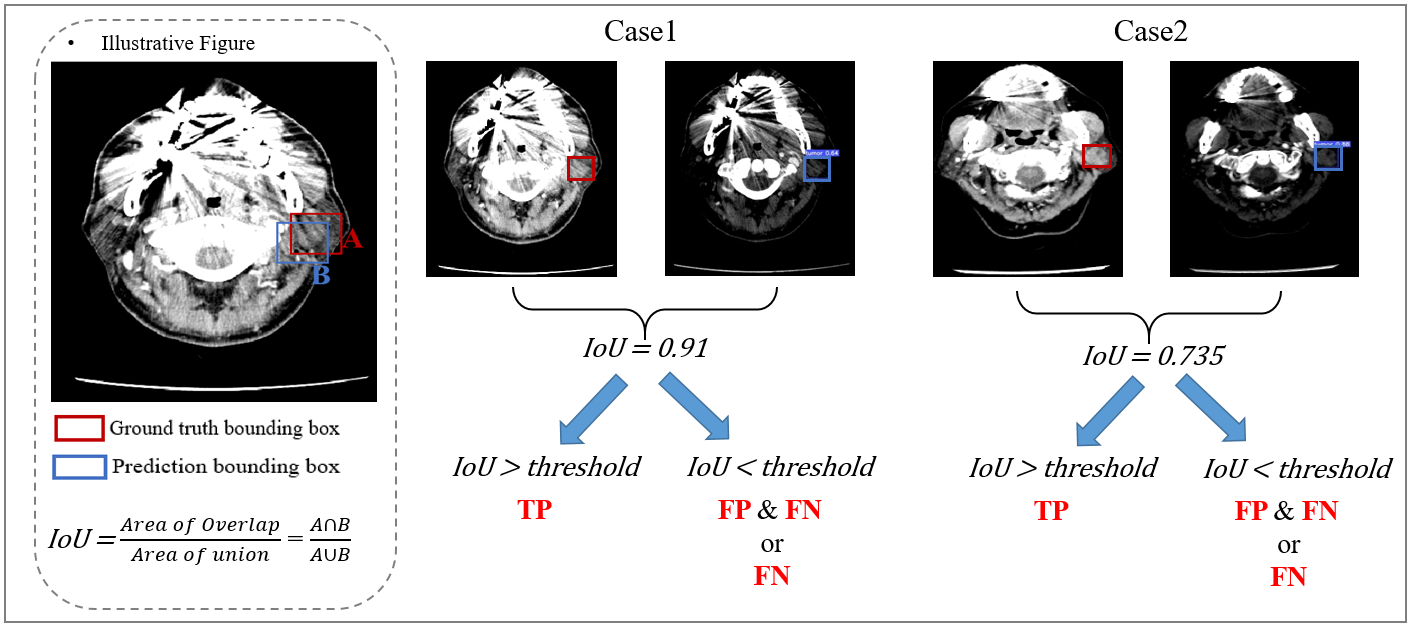
**Figure 5.**The evaluation mechanism in the slice space is demonstrated, presenting various prediction outcomes involving correct labels and predicted labels.

4.2 Based on the Detection Box Space(BD).

為解決前述問題，我們提出了一種基於偵測框空間的評估機制。我們將每張影像視為一個切片(Slice)。對於偵測框與腫瘤位置的交集比例(IoU)，若大於設定闕值，我們將其判定為真陽性(True Positive)；反之同時視為偽陽性和偽陰性。這包括兩種情況：(1) 應檢測出腫瘤卻未檢測到腫瘤，以及 (2) 不應檢測出腫瘤卻檢測出腫瘤。對於沒有腫瘤的檢測影像，若模型未檢測到腫瘤，我們視為真陰性(True Negative)；反之，若模型錯誤地檢測到腫瘤，則視為偽陽性(False Positive)。

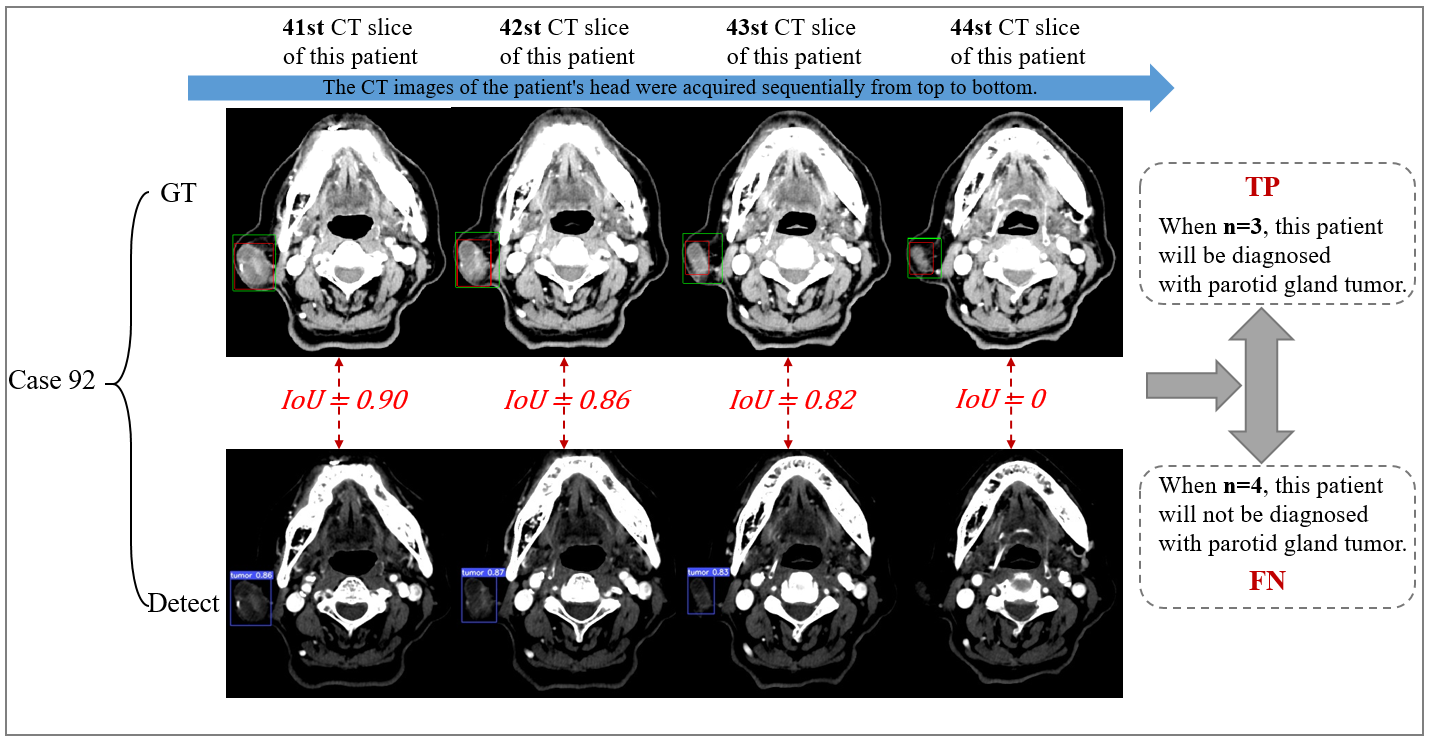
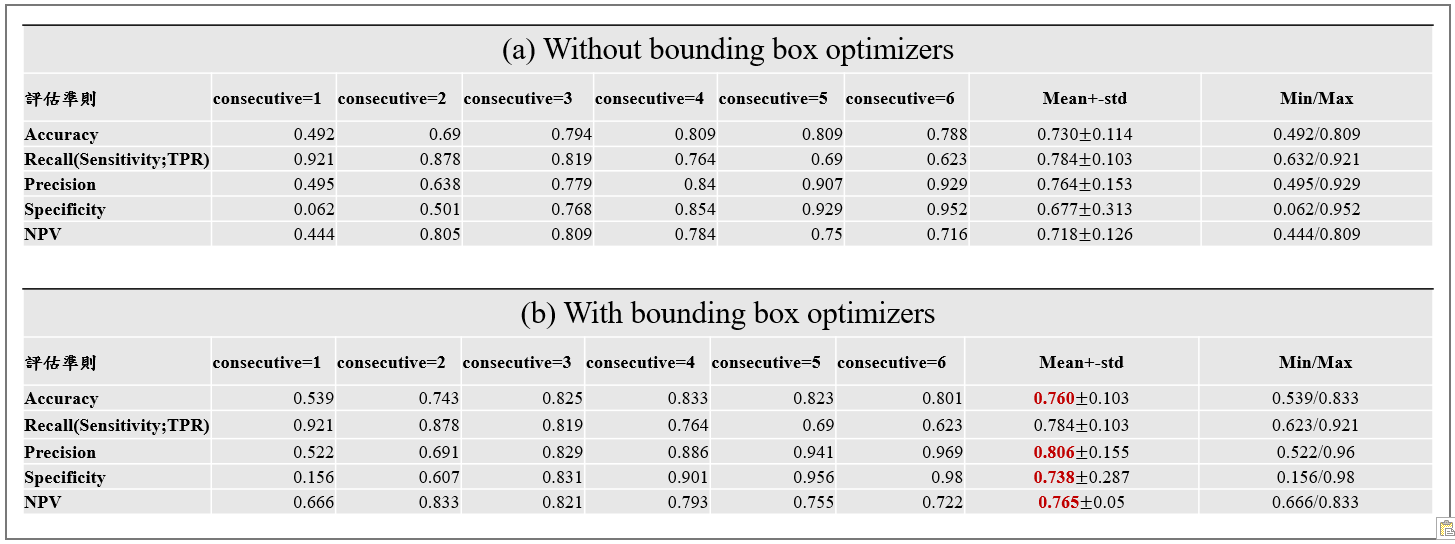
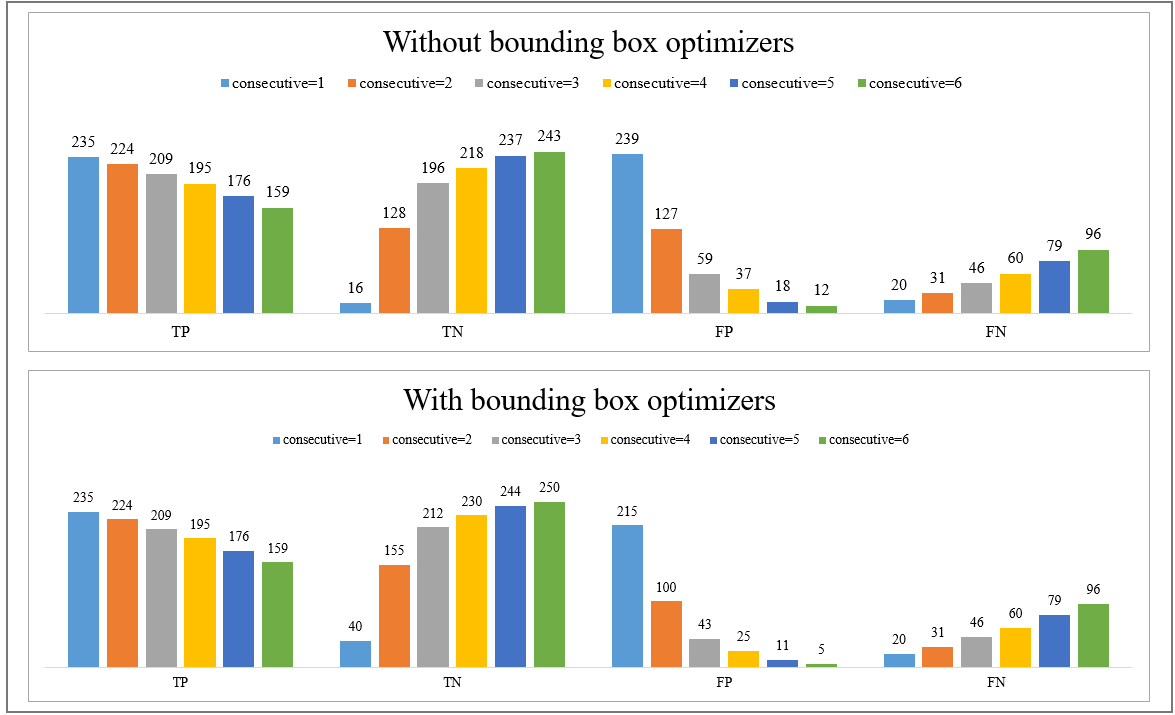
**Figure 6**. illustrates two scenarios regarding the Intersection over Union (IoU) between the detected bounding box and the actual tumor position: (a) simultaneous identification of false positives and false negatives when IoU is below a specified threshold, and (b) exclusive identification of false negatives. The variations in evaluation metrics with different IoU values are depicted.

為了緩解機制過於嚴格的問題，我們也引入了較為寬鬆的評估標準。當偵測框與腫瘤位置的IoU小於闕值時，統一將其視為偽陰性(False Negative)，不再計算偽陽性(圖3展示了在不同的評估機制下，不同IoU值對評估指標的變化情況)。這樣的改進有助於更合理地評估模型在腫瘤偵測方面的性能。

4.3 Based on the Tumor Space(BT).

**Figure 7.** presents two cases illustrating the model's predicted bounding boxes and corresponding ground truth labels, along with their respective IoU values

為提升臨床合理性，我們導入一種新的評估方式，考慮上下文給予的資訊，因為要判斷一位患者是否罹患腫瘤，儘管是醫生也常常需要整組影像進行上下文的關聯性，才能給予準確的判讀。我們以一個立體三維影像視為一個樣本，由於其具有連續性且上下文的密切相關性。在存在腫瘤的區域，我們要求連續的 n 張影像都需檢測出 IoU 大於闕值，視為存在腫瘤(True Positive)；反之視為未檢測出腫瘤(False Negative)。同樣地，在無腫瘤的區域，連續 n 張影像都未檢測出腫瘤才視為未存在腫瘤(True Negative)，反之視為存在腫瘤(False Positive)。這種方式更加符合臨床實際情況，考慮了連續影像的一致性，有助於提高對於腫瘤偵的可信度和臨床解釋性(圖5呈現了第92位病患在此評估標準的情況)。

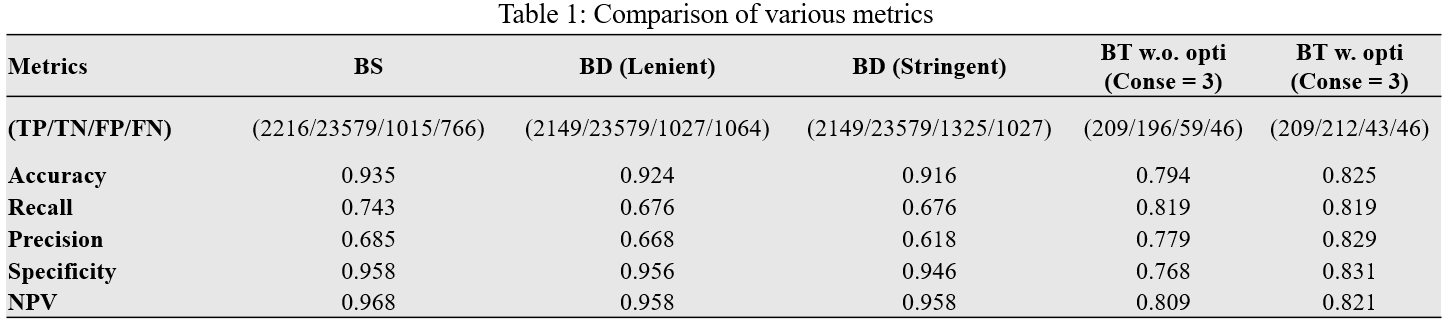
在取得CT影像時常常會因為資料難以取得，標記困難等等導致資料量不夠，因此在少量資料上進行訓練常常會有一些情況是訓練資料沒有的，這時在驗證集上如果出現這樣的情況，常常會有被模型誤檢測或者模型漏檢測的情況產生，為了避免這樣的情況下，我們提出了一種最佳化的方式，我們觀測患者的整組三維的CT影像設定一個立體的bounding box為我們需要檢測的區域，其餘的區域是不會出現腮腺腫瘤的我們便排除在此bounding box之外，這樣的方式可以有效避免那些不會出現腮腺腫瘤的區域偵測出腫瘤，這樣的方式我們稱之為Based on the Tumor with optimizer(BTWO)(圖6圖7比較了有無最佳化的結果)。

**Table 1.** presents the evaluation scores with and without optimization. On average, the optimized data tends to have higher scores.

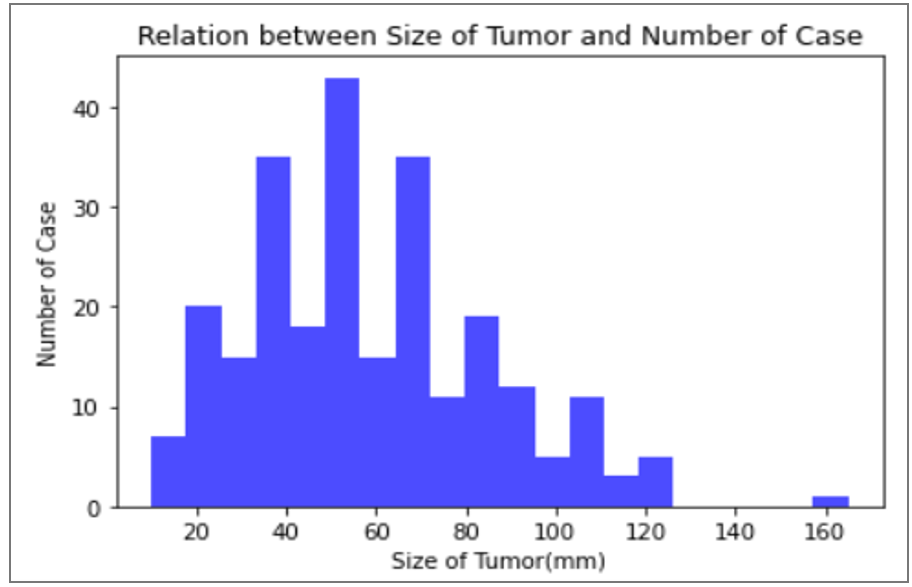
**Figure 8.** illustrates Case 92, where three consecutive tumor detections occur, highlighting variations in evaluation criteria under different standards.

**Figure 9.** directly illustrates the situations of TP, TN, FP, and FN, with Figure 6 derived from this dataset.

4.4 Experiment conclusion

根據4.1~4.3所述的方法，我們對YOLOv7在驗證集上的表現進行了全面評估。我們根據多項指標，包括Accuracy、Recall（敏感性; TPR）、Precision、Specificity、NPV，對四種方法的效果進行了評估（見表7）。我們觀察到，如果追求較高的準確率，則選擇BS作為評估標準；如果追求較高的敏感度，則選擇BT作為評估標準。在這樣的問題中，由於存在不同的需求，選擇方式也會有所不同。在醫學臨床上，我們通常希望最小化誤判的情況，因為誤判帶來的代價是巨大的，因此我們更傾向選擇敏感度較高的評估標準

**Table 2.** The table presents variations in evaluation metrics under different assessment methods. It is observed that the overall performance is optimized when utilizing the bounding box optimizer.

最終，我們選擇以基於腫瘤空間的評估方式作為我們的主要評估標準，IoU設定為0.001。一般的CT影像中，腫瘤呈現出一定的連續性。採用這種評估方式不僅能夠合理考慮上下文提供的資訊，以做出對患者是否罹患腮腺腫瘤的診斷決策，同時還能夠動態調整連續張數n觀察不同n值對評估指標的影響。圖8展示了在我們的驗證集上，患者腫瘤大小的分布情況。可以觀察到有一小部分患者的腫瘤大小低於20mm，即這些患者在各自的CT影像組中，出現腫瘤的影像不超過4張。換言之，如果將n設定為5，則這些患者將永遠不會被判定為患有腮腺腫瘤。對於n值的設定，我們建議根據資料分佈的實際情況進行調整。

**Figure 10.** illustrates the histogram of tumor sizes and patient counts. This can aid in determining the appropriate setting for the continuous slice number, n, based on the data distribution.

## 

**Figure 11.** displays the consecutive correct tumor positions and the detected tumor positions for three cases.

The 132nd patient

GT Object Detection

The 113nd patient

GT Object Detection

The 112nd patient

GT Object Detection

**Figure 12.** Sixteen examples of brain images with the related label provided by radiologists, designed to confirm the cancer localization

## 

## **5. Conclusion**

利用YOLOv7在腮腺腫瘤檢測方面取得了卓越的成果。透過綜合考量四種評估準則，我們能夠從不同的角度全面觀察模型的檢測表現。首先，從整體平均來看，經過最佳化的數據呈現出顯著的提升，這反映了模型在總體上的優越性。其次，根據各種不同的評估指標，例如Accuracy、Recall、Precision、Specificity等，我們能夠從不同的角度深入瞭解模型的表現。這種多角度的觀測使我們更全面地評估了腮腺腫瘤偵測的結果，有助於選擇最適合臨床需求的評估標準。總的來說，YOLOv7在腮腺腫瘤檢測上的卓越表現為醫學影像分析領域提供了一個可靠且高效的工具

**6. Reference**

[1] Emrah Gunduz  , Omer Faruk Alçin 2, Ahmet Kizilay  , Ismail Okan Yildirim.” [Deep learning model developed by multiparametric MRI in differential diagnosis of parotid gland tumors.](https://www.x-mol.com/paperRedirect/1528423765192486912)” [European Archives of Oto-Rhino-Laryngology](https://www.x-mol.com/paper/journal/2086?r_detail=1528423765192486912) ( IF 2.6 ) Pub Date : 2022-05-21 , DOI: [10.1007/s00405-022-07455-y](https://www.x-mol.com/paperRedirect/1528423765192486912)

[2] [Merve Önder](https://pubmed.ncbi.nlm.nih.gov/?term=%C3%96nder+M&cauthor_id=36832069) , [Cengiz Evli](https://pubmed.ncbi.nlm.nih.gov/?term=Evli+C&cauthor_id=36832069) , [Ezgi Türk](https://pubmed.ncbi.nlm.nih.gov/?term=T%C3%BCrk+E&cauthor_id=36832069) , [Orhan Kazan](https://pubmed.ncbi.nlm.nih.gov/?term=Kazan+O&cauthor_id=36832069) , [İbrahim Şevki Bayrakdar](https://pubmed.ncbi.nlm.nih.gov/?term=Bayrakdar+%C4%B0%C5%9E&cauthor_id=36832069) , [Özer Çelik](https://pubmed.ncbi.nlm.nih.gov/?term=%C3%87elik+%C3%96&cauthor_id=36832069) , [Andre Luiz Ferreira Costa](https://pubmed.ncbi.nlm.nih.gov/?term=Costa+ALF&cauthor_id=36832069) , [João Pedro Perez Gomes](https://pubmed.ncbi.nlm.nih.gov/?term=Gomes+JPP&cauthor_id=36832069) , [Celso Massahiro Ogawa](https://pubmed.ncbi.nlm.nih.gov/?term=Ogawa+CM&cauthor_id=36832069) , [Rohan Jagtap](https://pubmed.ncbi.nlm.nih.gov/?term=Jagtap+R&cauthor_id=36832069) , [Kaan Orhan](https://pubmed.ncbi.nlm.nih.gov/?term=Orhan+K&cauthor_id=36832069). “Deep-Learning-Based Automatic Segmentation of Parotid Gland on Computed Tomography Images” Diagnostics (Basel). 2023 Feb 4;13(4):581 doi: 10.3390/diagnostics13040581.

[3] [Xu Liu](https://pubmed.ncbi.nlm.nih.gov/?term=Liu%20X%5BAuthor%5D), MMed, [Yucheng Pan](https://pubmed.ncbi.nlm.nih.gov/?term=Pan%20Y%5BAuthor%5D), MD, [Xin Zhang](https://pubmed.ncbi.nlm.nih.gov/?term=Zhang%20X%5BAuthor%5D), MMed, [Yongfang Sha](https://pubmed.ncbi.nlm.nih.gov/?term=Sha%20Y%5BAuthor%5D), MMed,  [Shihui Wang](https://pubmed.ncbi.nlm.nih.gov/?term=Wang%20S%5BAuthor%5D), MS, [Hongzhe Li](https://pubmed.ncbi.nlm.nih.gov/?term=Li%20H%5BAuthor%5D), PhD, and [Jianping Liu](https://pubmed.ncbi.nlm.nih.gov/?term=Liu%20J%5BAuthor%5D), MD, PhD.

“A Deep Learning Model for Classification of Parotid Neoplasms Based on Multimodal Magnetic Resonance Image Sequences” [Laryngoscope.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10083903/) 2023 Feb; 133(2): 327–335. Published online 2022 May 16. doi: [10.1002/lary.30154](https://doi.org/10.1002%2Flary.30154)

# [4] [Hidetoshi Matsuo](https://pubmed.ncbi.nlm.nih.gov/?term=Matsuo+H&cauthor_id=33168936) , [Mizuho Nishio](https://pubmed.ncbi.nlm.nih.gov/?term=Nishio+M&cauthor_id=33168936) , [Tomonori Kanda](https://pubmed.ncbi.nlm.nih.gov/?term=Kanda+T&cauthor_id=33168936) , [Yasuyuki Kojita](https://pubmed.ncbi.nlm.nih.gov/?term=Kojita+Y&cauthor_id=33168936) , [Atsushi K Kono](https://pubmed.ncbi.nlm.nih.gov/?term=Kono+AK&cauthor_id=33168936) , [Masatoshi Hori](https://pubmed.ncbi.nlm.nih.gov/?term=Hori+M&cauthor_id=33168936) , [Masanori Teshima](https://pubmed.ncbi.nlm.nih.gov/?term=Teshima+M&cauthor_id=33168936) , [Naoki Otsuki](https://pubmed.ncbi.nlm.nih.gov/?term=Otsuki+N&cauthor_id=33168936) , [Ken-Ichi Nibu](https://pubmed.ncbi.nlm.nih.gov/?term=Nibu+KI&cauthor_id=33168936) , [Takamichi Murakami](https://pubmed.ncbi.nlm.nih.gov/?term=Murakami+T&cauthor_id=33168936). “Diagnostic accuracy of deep-learning with anomaly detection for a small amount of imbalanced data: discriminating malignant parotid tumors in MRI” Sci Rep. 2020 Nov 9;10(1):19388. doi: 10.1038/s41598-020-76389-4.

# [5] [Xianwu Xia](https://pubmed.ncbi.nlm.nih.gov/?term=Xia%20X%5BAuthor%5D),  [Bin Feng](https://pubmed.ncbi.nlm.nih.gov/?term=Feng%20B%5BAuthor%5D),  [Jiazhou Wang](https://pubmed.ncbi.nlm.nih.gov/?term=Wang%20J%5BAuthor%5D),  [Qianjin Hua](https://pubmed.ncbi.nlm.nih.gov/?term=Hua%20Q%5BAuthor%5D), [Yide Yang](https://pubmed.ncbi.nlm.nih.gov/?term=Yang%20Y%5BAuthor%5D), 4 [Liang Sheng](https://pubmed.ncbi.nlm.nih.gov/?term=Sheng%20L%5BAuthor%5D) [Yonghua Mou](https://pubmed.ncbi.nlm.nih.gov/?term=Mou%20Y%5BAuthor%5D), and [Weigang Hu](https://pubmed.ncbi.nlm.nih.gov/?term=Hu%20W%5BAuthor%5D). “Deep Learning for Differentiating Benign From Malignant Parotid Lesions on MR Images” [Front Oncol.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8262843/) 2021; 11: 632104.Published online 2021 Jun 23. doi: [10.3389/fonc.2021.632104](https://doi.org/10.3389%2Ffonc.2021.632104)

[6] [Ziyang Hu](https://pubmed.ncbi.nlm.nih.gov/?term=Hu%20Z%5BAuthor%5D),[Baixin Wang](https://pubmed.ncbi.nlm.nih.gov/?term=Wang%20B%5BAuthor%5D),  [Xiao Pan](https://pubmed.ncbi.nlm.nih.gov/?term=Pan%20X%5BAuthor%5D), 1[Dantong Cao](https://pubmed.ncbi.nlm.nih.gov/?term=Cao%20D%5BAuthor%5D),  [Antian Gao](https://pubmed.ncbi.nlm.nih.gov/?term=Gao%20A%5BAuthor%5D),  [Xudong Yang](https://pubmed.ncbi.nlm.nih.gov/?term=Yang%20X%5BAuthor%5D),[Ying Chen](https://pubmed.ncbi.nlm.nih.gov/?term=Chen%20Y%5BAuthor%5D), and [Zitong Lin](https://pubmed.ncbi.nlm.nih.gov/?term=Lin%20Z%5BAuthor%5D). “Using deep learning to distinguish malignant from benign parotid tumors on plain computed tomography images” [Front Oncol.](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9376440/) 2022; 12: 919088. Published online 2022 Aug 1. doi: [10.3389/fonc.2022.919088](https://doi.org/10.3389%2Ffonc.2022.919088)

[7] [Kubilay Muhammed Sunnetci](https://pubmed.ncbi.nlm.nih.gov/?term=Sunnetci+KM&cauthor_id=37271636) , [Esat Kaba](https://pubmed.ncbi.nlm.nih.gov/?term=Kaba+E&cauthor_id=37271636) , [Fatma Beyazal Celiker](https://pubmed.ncbi.nlm.nih.gov/?term=Celiker+FB&cauthor_id=37271636) , Ahmet Alkan. “Deep N etwork-Based Comprehensive Parotid Gland Tumor Detection” Acad Radiol. 2023 Jun 2:S1076-6332(23)00226-X. doi: 10.1016/j.acra.2023.04.028. Online ahead of print.

# [8] [Mohamed Nasor](https://dblp.org/pid/258/9510.html), Walid Obaid: “MRI Tumor Detection and Localization by Multiple Threshold Object Counting Technique. “[ICCA 2018](https://dblp.org/db/conf/icca3/icca2018.html#NasorO18): 158-161

# [9] [Prabira Kumar Sethy](https://ieeexplore.ieee.org/author/37085487539); [Baby Barnali Pattnaik](https://ieeexplore.ieee.org/author/37089650899); [Srikanta Dash](https://ieeexplore.ieee.org/author/37089650800). “Brain Tumour Localization using Image Processing Techniques” Published in: [2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)](https://ieeexplore.ieee.org/xpl/conhome/9986507/proceeding) DOI: [10.1109/I-SMAC55078.2022.9986502](https://doi.org/10.1109/I-SMAC55078.2022.9986502)

[10] [Burcu Selçuk](https://www.researchgate.net/profile/Burcu-Selcuk-2?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19).; [Tacha Serif](https://www.researchgate.net/profile/Tacha-Serif?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIn19). “Brain Tumor Detection and Localization with YOLOv8” Conference: 2023 8th International Conference on Computer Science and Engineering (UBMK) DOI:[10.1109/UBMK59864.2023.10286729](http://dx.doi.org/10.1109/UBMK59864.2023.10286729)

# [11] [Francesco Mercaldo](https://sciprofiles.com/profile/917645?utm_source=mdpi.com&utm_medium=website&utm_campaign=avatar_name), [Luca Brunese](https://sciprofiles.com/profile/1305789?utm_source=mdpi.com&utm_medium=website&utm_campaign=avatar_name), [Fabio Martinelli](https://sciprofiles.com/profile/172618?utm_source=mdpi.com&utm_medium=website&utm_campaign=avatar_name), [Antonella Santone](https://sciprofiles.com/profile/67586?utm_source=mdpi.com&utm_medium=website&utm_campaign=avatar_name) and [Mario Cesarelli](https://sciprofiles.com/profile/author/bFUzNWNFa0JSRFdqdlBDOVZjdGlwOEFkK0VOa0NsemVyalg5L3JKQ1VBMD0=?utm_source=mdpi.com&utm_medium=website&utm_campaign=avatar_name). “Object Detection for Brain Cancer Detection and Localization” Appl. Sci. 2023, 13(16), 9158; <https://doi.org/10.3390/app13169158>

[12] Jiang, P.; Ergu, D.; Liu, F.; Cai, Y.; Ma, B. “A Review of Yolo algorithm developments.” Procedia Comput. Sci. 2022, 199, 1066–1073.

[13] Chien-Yao Wang, [Alexey Bochkovskiy](https://dblp.org/pid/263/6975.html), [Hong-Yuan Mark Liao](https://dblp.org/pid/05/2757.html):  
“YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors”. [CVPR 2023](https://dblp.org/db/conf/cvpr/cvpr2023.html#WangBL23): 7464-7475

[14] Hurtik, P.; Molek, V.; Hula, J.; Vajgl, M.; Vlasanek, P.; Nejezchleba, T. Poly-“YOLO: Higher speed, more precise detection and instance segmentation for YOLOv3”. Neural Comput. Appl. 2022, 34, 8275–8290.

[15] Paul, S.; Ahad, D.M.T.; Hasan, M.M. “Brain Cancer Segmentation Using YOLOv5 Deep Neural Network”. arXiv 2022, arXiv:2212.13599.

[16] G. Jocher et al., “ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation,” Zenodo, Nov. 22, 2022. [Online].

[17] 　C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, “Scaled-YOLOv4: Scaling Cross Stage Partial Network,” arXiv:2011.08036, Feb. 2021. [Online].

[18] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” Apr. 2020. [Online]. Available: <https://arxiv.org/abs/2004.10934>.

[19] J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” arXiv:1804.02767, Apr. 2018. [Online]. Available: http://arxiv.org/abs/1804.02767. [Accessed: Aug. 09, 2023].

[20]　 J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” Dec. 2016. [Online] . doi: https://doi.org/10.48550/arxiv.1612.08242 [Accessed: Aug. 09, 2023].