

Introduction to deep learning

Midterm project report

YOLO for Medical Image Segmentation- Case study: Lung cancer

Group 9’s members:

|  |  |
| --- | --- |
| Vũ Đức Duy | 22BI13127 |
| Nguyễn Hoài Anh | 22BI13021 |
| Nguyễn Việt Anh | 22BI13032 |
| Phạm Hoàng Anh | 22BI13034 |
| Lâm Chí Cường | 22BI13068 |
| Nguyễn Át | 22BI13047 |

Contents

[**I.** **Introduction** 3](#_Toc178616074)

[**II.** **Dataset** 3](#_Toc178616075)

[**1.** **Data visualization** 3](#_Toc178616076)

[**2.** **Dataset description** 3](#_Toc178616077)

[**3.** **Example** 4](#_Toc178616078)

[**III.** **Model** 5](#_Toc178616079)

[**IV.** **Result** 6](#_Toc178616080)

[**V.** **Resources** 6](#_Toc178616081)

# **Introduction**

**You Only Look Once (YOLO)** is a real-time object detection model that is widely used in computer vision tasks and is known for its speed and accuracy in detecting objects within images and videos. YOLO proposes using an end-to-end neural network that makes predictions of bounding boxes and class probabilities all at once. It differs from the approach taken by previous object detection algorithms, which repurposed classifiers to perform detection.

In this project, we use **YOLO-v8** for **Medical Image Segmentation – Lung Cancer Nodules**. This report focuses on applying **YOLOv8**, the latest version of the series of real-time object detectors, offering cutting-edge performance in terms of accuracy and speed. We specifically investigate how well it can detect lung cancer nodules in medical images. In more detail, we use **YOLOv8-seg** and **YOLOv8-obb** for analysis so that we can provide some early results and better methods of treatment for patients.

# **Dataset**

## **Data visualization**

Our data has 343 images divided into 70% for training, 20% for validating and 10% for testing, each image size is 640 x 640 pixels. The data set is segmented into 2 tests in YOLOv8-seg and YOLOv8-obb and each test has 1 validation actual and prediction.

## **Dataset description**

-       x\_train contains 90% grayscale images.

-       y\_train corresponds to 90% labels denoting classes for each image in

x\_train

-       x\_test holds 10% grayscale images.

-       y\_test consists of 10% labels indicating classes for the images in x\_test.

## **Example**

In this test, we have run a small sample and here is the rusult and specific data for YOLOv8-seg and YOLOv8-obb:

|  |  |  |
| --- | --- | --- |
|  | Actual | Predicted |
| YOLOv8-seg |  |  |
| YOLOv8-obb |  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Size  (pixels) | mAPbox  50-95 | mAPmask  50-95 | Speed CPU  ONNX  (ms) | Speed A100  TensorRT  (ms) | params(M) | FLOPs(B) |
| YOLOv8x-seg | 640 | 53.4 | 43.4 | 712.1 | 4.02 | 71.8 | 344.1 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Model | Size  (pixels) | mAPbox  50-95 | mAPmask  50-95 | Speed CPU  ONNX  (ms) | Speed A100  TensorRT  (ms) | params(M) | FLOPs(B) |
| YOLOv8x-obb | 640 | 53.4 | 43.4 | 712.1 | 4.02 | 71.8 | 344.1 |

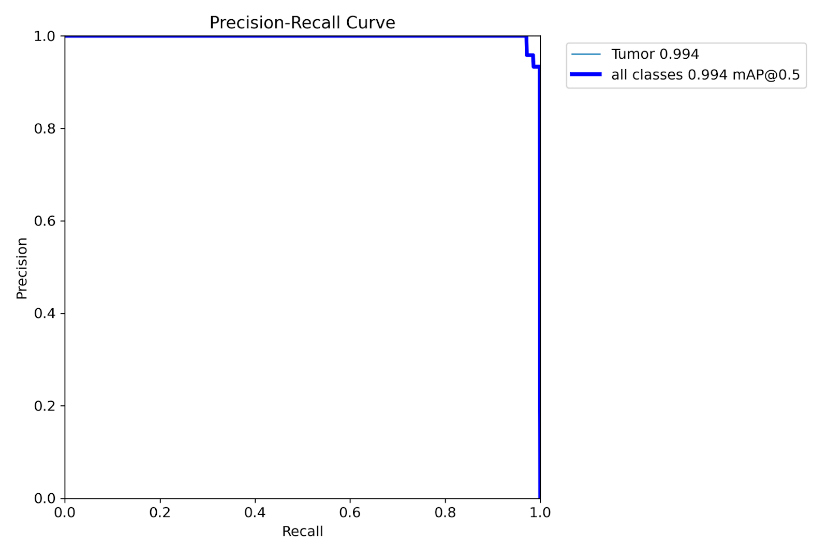
# **Model**

The design of YOLOv8 is based on earlier versions of YOLO algorithms. YOLOv8 makes use of a convolutional neural network (CNN) that is comprised of two main components: the backbone and the head. The foundation of YOLOv8 is based on an altered iteration of the CSPDarknet53 design. This structure includes 53 convolutional layers and utilizes partial crossover phase connections to enhance communication between the various layers. YOLOv8's main component is made up of numerous convolutional layers and is followed by several fully connected layers. These layers predict bounding boxes, objectness score and class probabilities for detected objects in an image. A crucial aspect of YOLOv8 is the incorporation of a self-attention mechanism in the network's head. This mechanism enables the model to concentrate on various sections of the image and modify the significance of different characteristics depending on their relevance to the task. Another significant aspect of YOLOv8 is its capability to conduct object detection at multiple scales. The model uses a feature pyramid network to identify objects of various sizes and scales in an image. This feature pyramid network is made up of several layers that identify objects of various sizes, enabling the model to recognize both big and small objects in an image.

# **Result**

* **YOLOv8-obb**

|  |  |  |
| --- | --- | --- |
|  |  | |
|  |  | |
|  | |  |



* **The bar chart image shows several plots:**

A bar chart labeled "Tumor" with around 250 instances. Two scatter plots showing the distribution of data points. One scatter plot shows width vs. height, while the other shows x vs. y coordinates

* **F1-Confidence Curve:**

Shows the relationship between F1 score and confidence. The curve reaches a peak F1 score of about 0.98 at a confidence level of 0.751. Performance is generally high across most confidence levels

* **Recall-Confidence Curve:**

Demonstrates the relationship between recall and confidence. The curve maintains perfect recall (1.0) up to a high confidence level before sharply declining

* **Precision-Confidence Curve:**

Illustrates how precision changes with confidence. Precision increases steadily with confidence, reaching 1.0 at a confidence of 0.752.

* **Precision-Recall Curve:**

Shows the trade-off between precision and recall. The curve indicates very high performance, with both precision and recall close to 1.0. The area under the curve (AUC) or mean Average Precision (mAP) at 0.5 is 0.994.

* **YOLOv8-seg**

|  |  |
| --- | --- |
|  |  |
|  |  |

* **The bar chart image shows several plots:**

The bar chart shows scatter plots and a bar chart. The bar chart indicates there are around 250 instances labeled as "Tumor". The scatter plots show the distribution of data points based on x, y coordinates and width/height.

* **F1-Confidence Curve**

Shows the balances precision and recall, peaks at around 0.91 at a confidence threshold of 0.326.

* **Recall-Confidence Curve**

Demonstrates a Recall-Confidence Curve. It shows how the recall of the model changes with different confidence thresholds. The curve indicates high recall at low confidence levels, which gradually decreases as confidence increases.

* **Precision-Confidence Curve**

Illustrates a Precision-Confidence Curve. The precision of the model increases with higher confidence thresholds, reaching 1.0 at a confidence of 0.678.

* **Precision-Recall Curve:**

The last image is a Precision-Recall Curve. It illustrates the trade-off between precision and recall as the model's decision threshold is varied. The curve shows high precision across most recall values, with an area under the curve (mAP@0.5) of 0.937.

# **Resources**

1. <https://github.com/ultralytics/yolov5>
2. <https://www.kaggle.com/code/younesselbrag/yolov5-on-costume-dataset-lung-cancer-nodules-dete/input>
3. <https://universe.roboflow.com/kfupm-7hvnm/lung-tumor-2xcgn/dataset/3?fbclid=IwY2xjawFm90xleHRuA2FlbQIxMAABHYwwOtgz_DjNWOFzzugF4XLkM18bBUsU7up4vf6suhC79WiRC-TIs4lYOg_aem_Wd0kkIv-VnPttCsEvbt8ng>
4. [https://docs.ultralytics.com/models/yolov8/](https://docs.ultralytics.com/models/yolov8/?fbclid=IwZXh0bgNhZW0CMTAAAR3-cG5rG0dwBIe09d79J1ZshvHuhGmj2jazHYw-ZAl7gTeiICIs7rvzL9Q_aem_qxm8eZ37JgTHzs2pBlrv4Q)