

# PHÂN ĐOẠN ẢNH Y KHOA TỰ ĐỘNG DỰA TRÊN YOLO-NAS VÀ MEDSAM2 KẾT HỢP NÉN MÔ HÌNH SAU HUẤN LUYỆN

AUTOMATED MEDICAL IMAGE SEGMENTATION VIA YOLO-NAS AND MEDSAM2  
WITH POST-TRAINING COMPRESSION

Nguyễn Ngọc Lan Thi - 240101077

# Tóm tắt

- Lớp: CS2205.FEB2025
- Link Github: <https://github.com/thinguyen243/CS2205.FEB2025>
- Link YouTube video: <https://youtu.be/gRA6sV6vmoU>



Họ và tên: Nguyễn Ngọc Lan Thi  
MSHV: 240101077

# Giới thiệu

Phân đoạn ảnh y khoa là bước quan trọng hỗ trợ chẩn đoán  
MedSAM2 chính xác cao nhưng lệ thuộc vào **prompt thủ công**  
Mô hình hiện hành **có kích thước lớn** → khó triển khai thực tế

## Giải pháp:

- ❑ **Tinh chỉnh YOLO-NAS** để sinh bounding box dùng cho **MedSAM2**
- ❑ Dùng kỹ thuật **Quantization** để nén mô hình sau khi huấn luyện

**Dữ liệu:** 20.000 ảnh y khoa và 1.000 video y khoa từ NIH ChestX-ray14, BraTS, Decathlon  
**So sánh:** FastSAM, MobileSAM, nnU-Net

# Giới thiệu

## Câu hỏi nghiên cứu:

1. Có thể kết hợp YOLO-NAS và MedSAM2 để tạo pipeline phân đoạn ảnh y khoa hoàn toàn tự động, thay thế cho prompt thủ công không?
2. Bounding box do YOLO-NAS sinh ra có đáng tin cậy không?
3. Việc áp dụng Quantization ảnh hưởng thế nào đến độ chính xác và tốc độ của hệ thống phân đoạn?

# Mục tiêu

- ❑ Xây dựng pipeline tự động phát hiện và phân đoạn ảnh y khoa được kết hợp bởi **YOLO-NAS** và **MedSAM2**
- ❑ Đánh giá được độ chính xác của bounding box tự sinh từ YOLO-NAS
- ❑ Nén mô hình bằng **Quantization** nhưng vẫn giữ hiệu năng cao

# Nội dung và Phương pháp

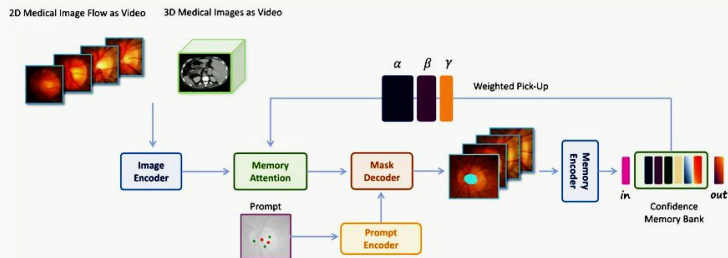
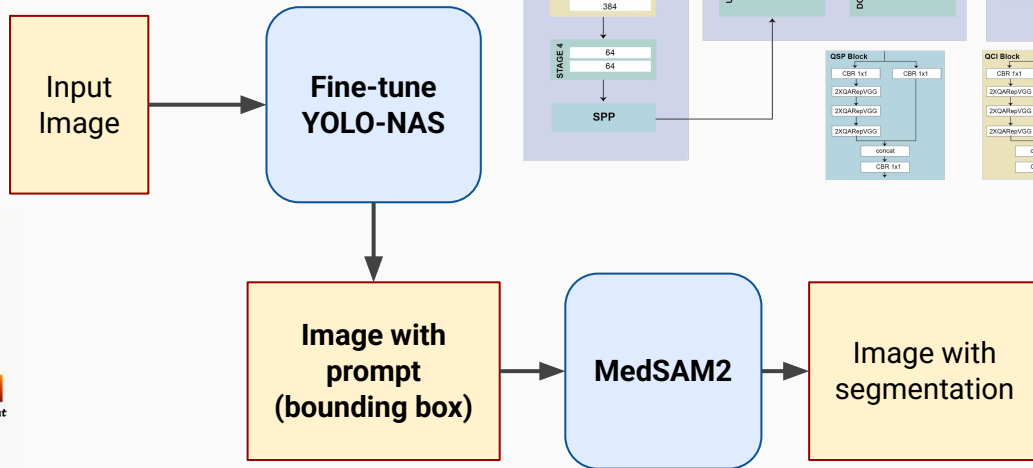
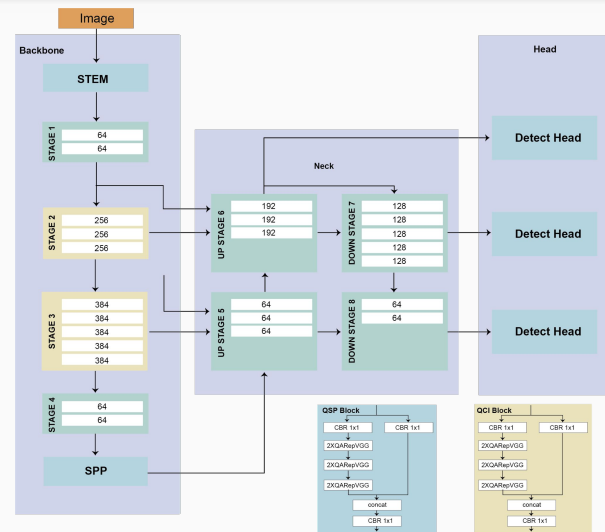
## 1. Thu thập và chuẩn bị dữ liệu:

20.000 ảnh và 1.000 video, chuẩn hóa về 640x640, split 70 train/20 val/10 test

## 3. Phân đoạn với MedSAM2: Sử dụng bounding box từ Fine-tune YOLO-NAS để phân đoạn

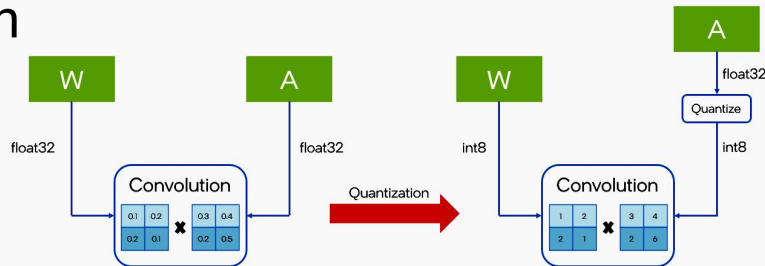
## 2. Huấn luyện YOLO-NAS:

Tinh chỉnh với dữ liệu y khoa để sinh bounding box tự động



# Nội dung và Phương pháp

**Nén mô hình:** Dùng kỹ thuật Quantization  
(FLOAT32  $\rightarrow$  INT8)



**Đánh giá và so sánh:** So với FastSAM, MobileSAM, nnU-Net

❑ IoU, Dice Score

$$IOU(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad dice(A, B) = \frac{2|A \cap B|}{|A| + |B|}$$

❑ Tốc độ xử lý (ms/ảnh)

❑ Kích thước mô hình

# Kết quả dự kiến

Mô hình phân đoạn ảnh y khoa **tự động**:

- Độ chính xác cao: **trên 75% theo IoU** và **trên 85% theo Dice Score**
- Tốc độ suy luận nhanh: **dưới 500ms** mỗi ảnh
- Kích thước mô hình nhỏ: **không quá 100MB**

Mô hình	IoU/Dice	Kích thước	Prompt thủ công	Tốc độ
<b>YOLO-NAS + MedSAM2 + Quantization</b>	>0.75/>0.85	$\leq 100\text{MB}$	Không	< 500ms
FastSAM	0.65 / 0.72	$\sim 120\text{MB}$	Không	Trung bình
MobileSAM	0.60 / 0.68	$\sim 90\text{MB}$	Không	Nhanh
nnUNet	0.78 / 0.83	> 200MB	Có	Chậm



# Tài liệu tham khảo

- [1]. Jiayuan Zhu et al.: Medical SAM 2: Segment medical images as video via Segment Anything Model 2. arXiv:2408.00874, 2024
- [2]. Juan Terven, Diana Cordova-Esparza: A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. arXiv:2304.00501, 2023
- [3]. Isensee, F. et al.: nnU-Net Revisited: A Call for Rigorous Validation in 3D Medical Image Segmentation. MICCAI, 2024.
- [4]. Victor Ion Butoi et al.: UniverSeg: Universal Medical Image Segmentation. ICCV, 2023, pp. 21438-21451
- [5]. R. Azad et al.: Medical Image Segmentation Review: The Success of U-Net. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 12, 2024, pp. 10076-10095
- [6]. Chengtao Lv, Hong Chen, Jinyang Guo, Yifu Ding, Xianglong Liu: PTQ4SAM: Post-Training Quantization for Segment Anything. CVPR, 2024, pp. 15941-15951