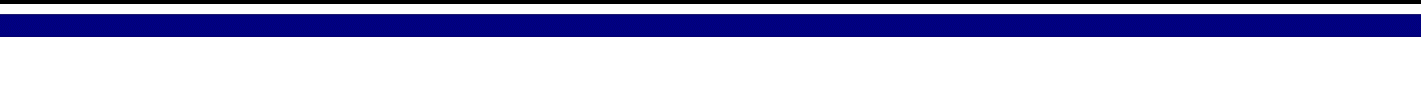
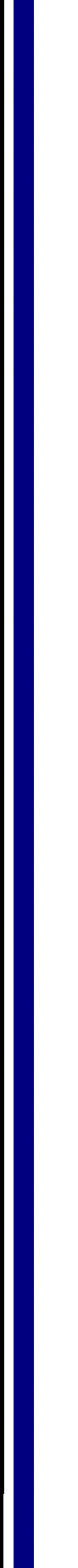
A blue and black rectangles

Description automatically generated



BỘ GIÁO DỤC VÀ ĐÀO TẠOA blue and red text on a white background

Description automatically generated

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ TP. HCM**

BÁO CÁO ĐỒ ÁN MÔN TRÍ TUỆ NHÂN TẠO

# **Human Detection in Thermal Imaging Using YOLO**

Ngành: **Công Nghệ Thông Tin**

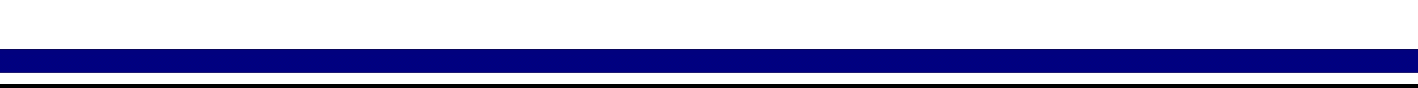
Tên học phần: **Trí Tuệ Nhân Tạo**

Giảng viên hướng dẫn: Huỳnh Quốc Bảo

Sinh viên thực hiện đồ án:

|  |  |  |
| --- | --- | --- |
| Trương Cao Vũ | MSSV: 2280603763 | Lớp: 22DTHG4 |
| Nguyễn Hòa Nam | MSSV: 2280602008 | Lớp: 22DTHG4 |
| Bùi Khôi Nguyên | MSSV: 2280602103 | Lớp:22DTHG4 |
|  |  |  |

TP.HCM, Tháng 9 Năm 2024

A blue and black rectangles

Description automatically generated

1

**LỜI MỞ ĐẦU**

Nhận diện con người trong ảnh nhiệt là một lĩnh vực nghiên cứu quan trọng, đặc biệt trong các ứng dụng an ninh, cứu hộ và giám sát. Sử dụng các công nghệ như YOLO (You Only Look Once), một trong những mô hình học sâu tiên tiến, chúng ta có thể phát hiện con người một cách hiệu quả và nhanh chóng trong các điều kiện ánh sáng yếu, chẳng hạn như ban đêm. Mục tiêu của nghiên cứu này là phát triển và cải thiện khả năng nhận diện người từ các hình ảnh nhiệt, cung cấp giải pháp tối ưu cho các tình huống khẩn cấp và tăng cường an ninh. Bài viết sẽ trình bày những kết quả đạt được từ việc áp dụng mô hình YOLO vào bài toán nhận diện người trong ảnh nhiệt, cùng với những thách thức và cơ hội trong lĩnh vực này.

**MỤC LỤC**

[**MỤC LỤC 3**](#_Toc155484912)

[**MỤC LỤC HÌNH ẢNH 5**](#_Toc155484913)

[**CHƯƠNG 1: GIỚI THIỆU 6**](#_Toc155484914)

**1.1** [**Tổng quan về nhận diện người (Human Detection) trong ảnh nhiệt. 6**](#_Toc155484915)

[**1.2 Nhiệm vụ đồ án 6**](#_Toc155484916)

[**1.2.1 Xây dựng hệ thống nhận diện người trong ảnh nhiệt sử dụng YOLO 6**](#_Toc155484916)

[**1.3 Ý nghĩa và ứng dụng của nhận diện người trong an ninh, cứu hộ , giám sát 7**](#_Toc155484917)

[**CHƯƠNG 2: CƠ SỞ LÝ THUYẾT 10**](#_Toc155484920)

[**2.1 Ảnh nhiệt 10**](#_Toc155484921)

[**2.1.1 Định nghĩa ảnh nhiệt và đặc điểm của nó 10**](#_Toc155484922)

[**2.1.2 Mục tiêu 10**](#_Toc155484923)

[**2.2 YOLO (You Only Look Once) 10**](#_Toc155484925)

[**2.2.1 Nguyên lý hoạt động của YOLO 10**](#_Toc155484926)

[**2.2.2 Các phiên bản của YOLO 11**](#_Toc155484927)

[**2.3 Phương pháp phát hiện đối tượng trong ảnh 20**](#_Toc155484935)

[**2.3.1 Các phương pháp truyền thống 20**](#_Toc155484936)

[**2.3.2 So sánh với YOLO 21**](#_Toc155484937)

[**CHƯƠNG 3: KẾT QUẢ THỰC NGHIỆM 27**](#_Toc155484942)

[**3.1 Bộ dữ liệu ảnh nhiệt 27**](#_Toc155484943)

[**3.1.1. Tóm tắt dữ liệu 27**](#_Toc155484944)

[**3.1.2 Mô tả bộ dữ liệu sử dụng 28**](#_Toc155484945)

[**3.2 Xử lí và chuẩn bị dữ liệu 38**](#_Toc155484948)

[**3.3 Kết quả của mô hình YOLO 47**](#_Toc155484952)

[**3.1.2 Đánh giá hiệu năng mô hình trên ảnh nhiệt 28**](#_Toc155484945)

[**3.1.3 Phân tích kết quả và độ chính xác của mô hình 28**](#_Toc155484945)

[**CHƯƠNG 4: KẾT LUẬN VÀ KIẾN NGHỊ 51**](#_Toc155484953)

[**4.1 Kết luận hiệu quả của mô hình YOLO trong nhận diện người từ ảnh nhiệt 51**](#_Toc155484954)

[**4.2 Hướng phát triển đề xuất cải tiến và các ứng dụng tương lai 52**](#_Toc155484955)

**MỤC LỤC HÌNH ẢNH**

**CHƯƠNG 1: GIỚI THIỆU**

**1.1 Tổng quan về nhận diện người (Human Detection) trong ảnh nhiệt:**

- Nhận diện người trong ảnh nhiệt là một lĩnh vực quan trọng trong công nghệ nhận diện đối tượng, được sử dụng rộng rãi trong các ứng dụng an ninh, cứu hộ, và giám sát. Ảnh nhiệt sử dụng bức xạ hồng ngoại phát ra từ các đối tượng để xác định vị trí và hình dáng của chúng, từ đó hỗ trợ trong việc phát hiện và phân loại người trong các điều kiện ánh sáng yếu hoặc trong môi trường khắc nghiệt.

- Sự phát triển của công nghệ máy học và trí tuệ nhân tạo đã mở ra nhiều hướng đi mới cho việc cải thiện độ chính xác và hiệu quả trong nhận diện người qua ảnh nhiệt.

**1.2 Nhiệm vụ đồ án**

### **1.2.1** **Xây dựng hệ thống nhận diện người trong ảnh nhiệt sử dụng YOLO:**

- Đồ án này đặt mục tiêu phát triển một hệ thống nhận diện người hiệu quả bằng cách áp dụng mô hình YOLO (You Only Look Once) vào ảnh nhiệt. YOLO là một trong những phương pháp tiên tiến nhất hiện nay trong việc phát hiện đối tượng, cho phép xử lý nhanh chóng và chính xác thông tin hình ảnh.

## **1.3** Ý nghĩa và ứng dụng của nhận diện người trong an ninh, cứu hộ, và giám sát

***1.3.1 Ý nghĩa***

- Nhận diện người trong ảnh nhiệt mang lại nhiều ý nghĩa quan trọng trong xã hội hiện đại. Công nghệ này không chỉ giúp cải thiện khả năng an ninh và giám sát mà còn có thể góp phần vào việc cứu hộ và bảo vệ tính mạng con người. Việc phát hiện nhanh chóng các cá nhân trong điều kiện ánh sáng yếu giúp tăng cường an ninh cho các khu vực công cộng, hỗ trợ lực lượng cứu hộ trong các tình huống khẩn cấp, và nâng cao nhận thức cộng đồng về an toàn.

***1.3.2 Ứng dụng***

Ứng dụng của nhận diện người trong ảnh nhiệt rất đa dạng, bao gồm nhưng không giới hạn ở:

1. **An ninh**: Giám sát và phát hiện xâm nhập trái phép trong các khu vực nhạy cảm.
2. **Cứu hộ**: Hỗ trợ lực lượng cứu hộ trong việc xác định vị trí người bị nạn trong tình huống hỏa hoạn hoặc thiên tai.
3. **Giám sát giao thông**: Phát hiện và theo dõi các hành vi bất thường của người tham gia giao thông.
4. **Phân tích hành vi**: Nghiên cứu hành vi con người trong các không gian công cộng nhằm cải thiện trải nghiệm người dung

Việc áp dụng công nghệ nhận diện người trong ảnh nhiệt có thể mang lại lợi ích thiết thực cho nhiều lĩnh vực trong cuộc sống hàng ngày.

**CHƯƠNG 2: CƠ SỞ LÝ THUYẾT**

**2.1 Ảnh nhiệt :**

**2.1.1 Định nghĩa ảnh nhiệt và đặc điểm của nó :**

### **2.1.1.1 Định nghĩa:**

### - Ảnh nhiệt là hình ảnh thể hiện phân bố nhiệt độ của các vật thể thông qua bức xạ hồng ngoại mà chúng phát ra. Mỗi vật thể có nhiệt độ khác nhau sẽ phát ra mức bức xạ khác nhau, cho phép chúng ta nhận diện và phân tích chúng dựa trên nhiệt độ.



Hình ảnh 2.1.1.1: Ảnh nhiệt

### **2.1.1.2 Đặc điểm :**

- Các đặc điểm chính của ảnh nhiệt bao gồm:

1. **Khả năng phát hiện**: Ảnh nhiệt có thể phát hiện người và vật thể trong điều kiện ánh sáng yếu hoặc bóng tối.
2. **Chi tiết**: Cung cấp thông tin chi tiết về nhiệt độ và sự phân bố nhiệt độ trên bề mặt vật thể.

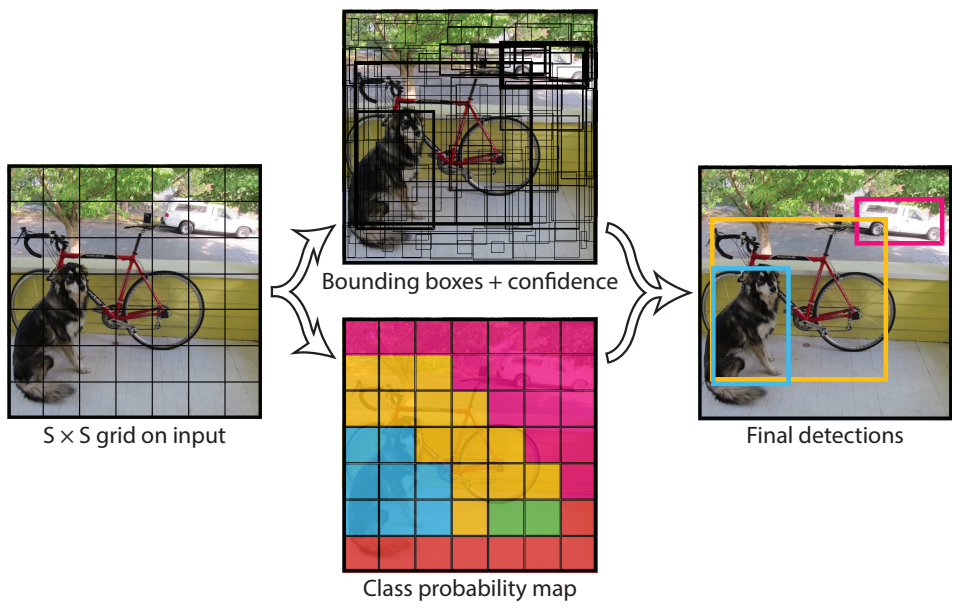
**2.1.2 Mục tiêu**

Mục tiêu của việc sử dụng ảnh nhiệt trong nhận diện người là cải thiện độ chính xác và tốc độ trong việc phát hiện con người, đặc biệt trong các tình huống khẩn cấp như cứu hộ hoặc giám sát an ninh.

**2.2 YOLO (You Only Look Once):**

**2.2.1 Nguyên lí hoạt động của YOLO:**

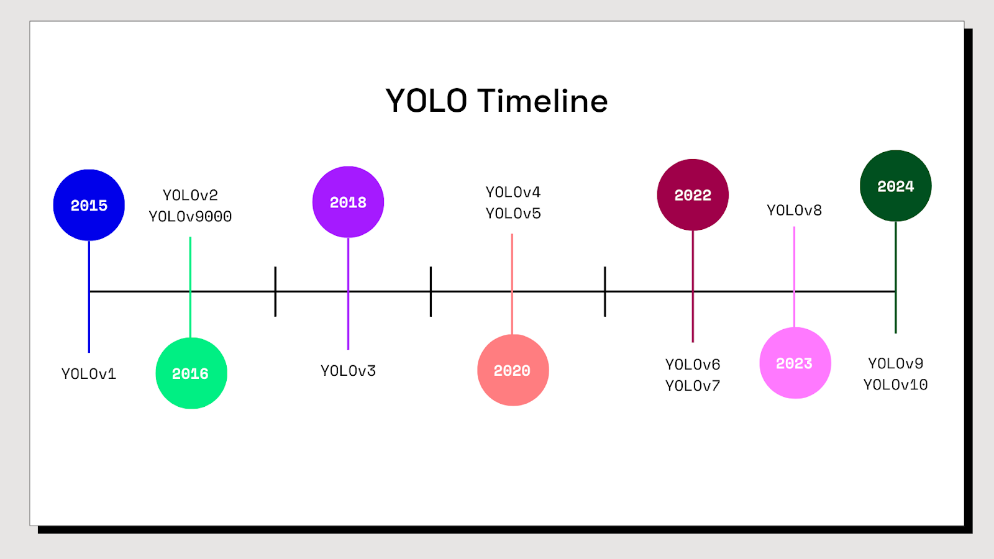
- YOLO là một thuật toán phát hiện đối tượng nhanh chóng, hoạt động theo nguyên lý chia hình ảnh thành lưới và dự đoán các bounding box cùng với xác suất cho từng lớp đối tượng trong mỗi ô lưới. Việc này giúp giảm thời gian xử lý và tăng hiệu quả phát hiện.



Hình 2.2.1: Cách YOLO hoạt động

**2.2.2 Các phiên bản của YOLO:**

- YOLO đã trải qua nhiều phiên bản từ YOLOv1 đến YOLOv11, mỗi phiên bản cải tiến về độ chính xác và hiệu suất. Các phiên bản mới nhất, như YOLOv11, đã cho thấy khả năng phát hiện tốt hơn trong môi trường phức tạp.

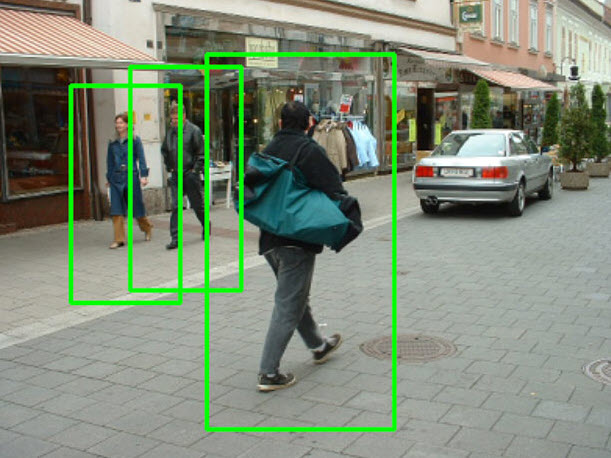


Hình 2.2. 2 : Phiên bản YOLO qua các năm

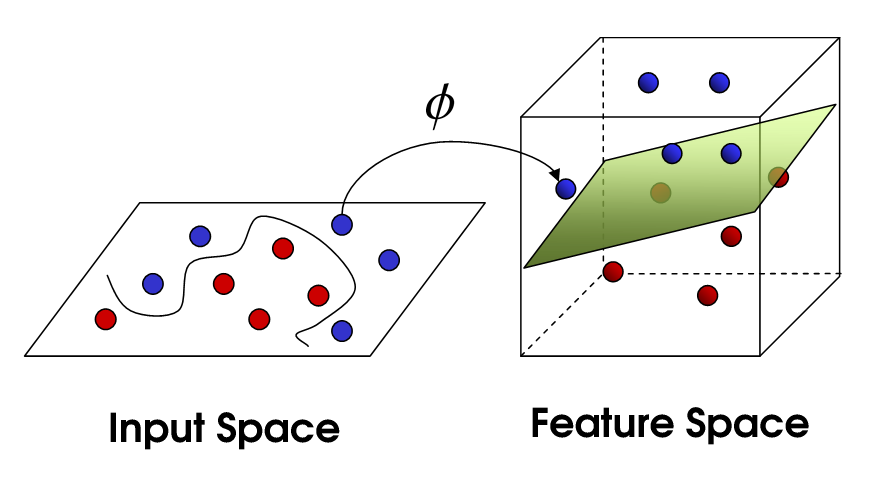
**2.2.3 Phương pháp phát hiện đối tượng trong ảnh:**

**2.3.1** **Các phương pháp truyền thống:**

Trước khi có YOLO, các phương pháp truyền thống như HOG (Histogram of Oriented Gradients) và SVM (Support Vector Machine) thường được sử dụng để phát hiện đối tượng, tuy nhiên chúng thường chậm hơn và kém chính xác hơn trong các tình huống thực tế.



Hình 2.2. 1.1Phương pháp truyền thống HOG



Hình 2.2.3.2: Phương pháp SVM

**2.3.2 So sánh với YOLO**

YOLO vượt trội hơn nhờ khả năng phát hiện đồng thời nhiều đối tượng trong một khung hình mà không cần phân tích từng phần riêng biệt. Điều này làm cho YOLO trở thành một lựa chọn lý tưởng cho các ứng dụng yêu cầu độ chính xác cao và tốc độ nhanh.



Hình 2.3.2: So sánh YOLO với các ứng dụng khác

**CHƯƠNG 3: KẾT QUẢ THỰC NGHIỆM**

**3.1 Bộ dữ liệu Ảnh nhiệt :**

**3.1.1 Tóm tắt dữ liệu:**

- Bộ dữ liệu ảnh nhiệt được sử dụng trong nghiên cứu này được cung cấp bởi Roboflow, bao gồm các hình ảnh thu thập từ nhiều nguồn nhằm phản ánh nhiều điều kiện môi trường và ánh sáng khác nhau. Bộ dữ liệu này đa dạng và chứa các đối tượng người ở nhiều bối cảnh và điều kiện nhiệt độ khác nhau, giúp mô hình học các đặc điểm quan trọng để phát hiện người trong các tình huống khác nhau.

**Dữ liệu có thể được tải từ liên kết sau: [Person Detection Dataset on Roboflow](https://universe.roboflow.com/smart2/persondection-61bc2/dataset/5#).**

**3.1.2 Xử lý và chuẩn bị dữ liệu:**

- **Phân chia dữ liệu:** Bộ dữ liệu được chia thành ba phần: 8000 ảnh cho tập huấn luyện, 1000 ảnh cho tập validation và 1000 ảnh cho tập test nhằm đảm bảo tính khách quan khi đánh giá hiệu suất mô hình. (Phần chia này được xử lý sẵn trên Roboflow).

- **Chuẩn hóa kích thước:** Tất cả ảnh được chuẩn hóa về **kích thước 640x640 pixel** để tối ưu khả năng xử lý của mô hình YOLOv8.

- **Augmentation:**

+, Sau khi training **15 epochs - model v8n(nano) - No Augmentation,** (điều này cho phép đánh giá chính xác khả năng phát hiện của mô hình YOLOv8 ban đầu mà không có sự điều chỉnh về dữ liệu).

+, Ta chuyển **sang training 50 epochs – model v5m(medium) with Augmentation** để tăng cường dữ liệu, giúp cải thiện khả năng tổng quát của mô hình và giảm thiểu hiện tượng overfitting.

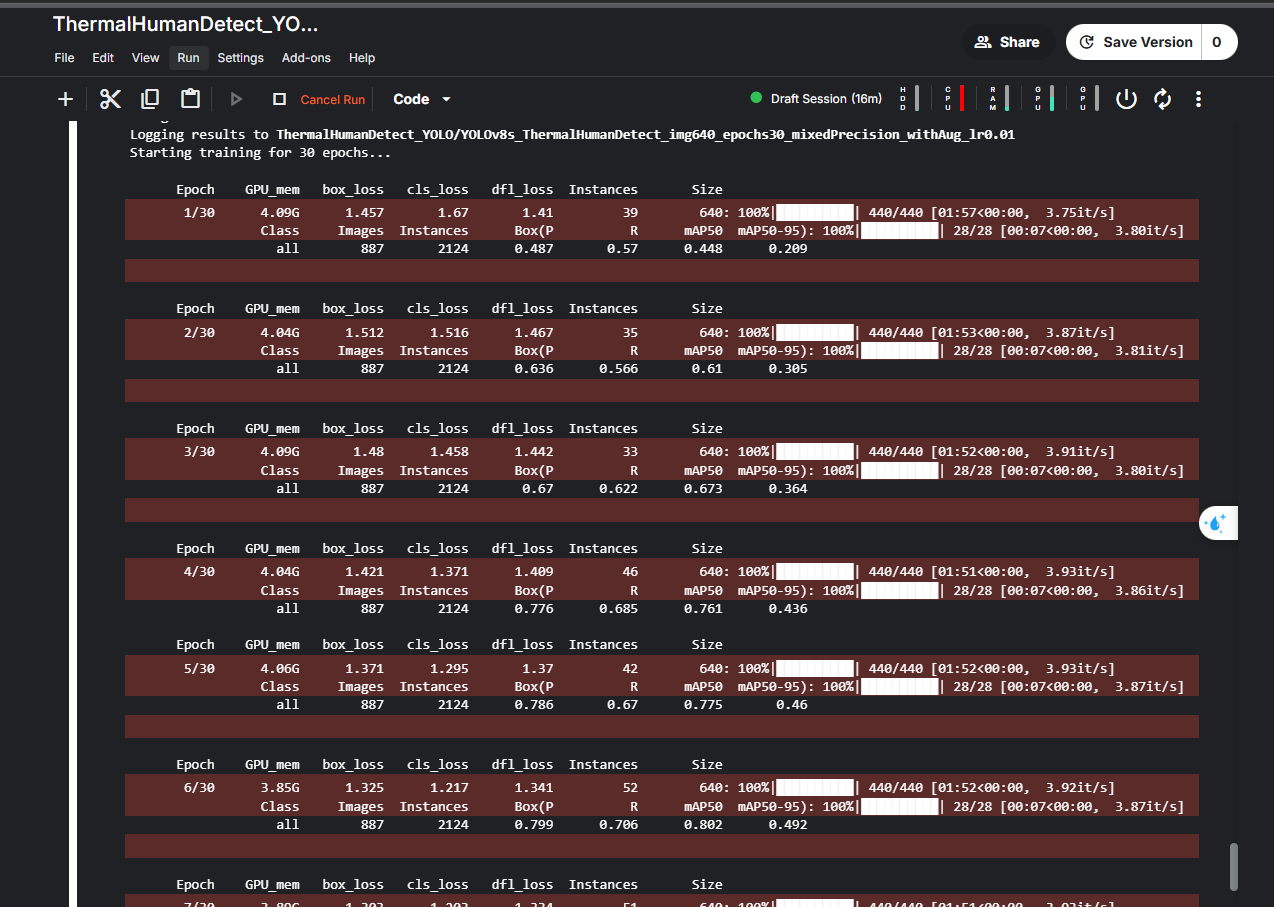
### **3.1.3 Quá trình training model YOLOv8:**

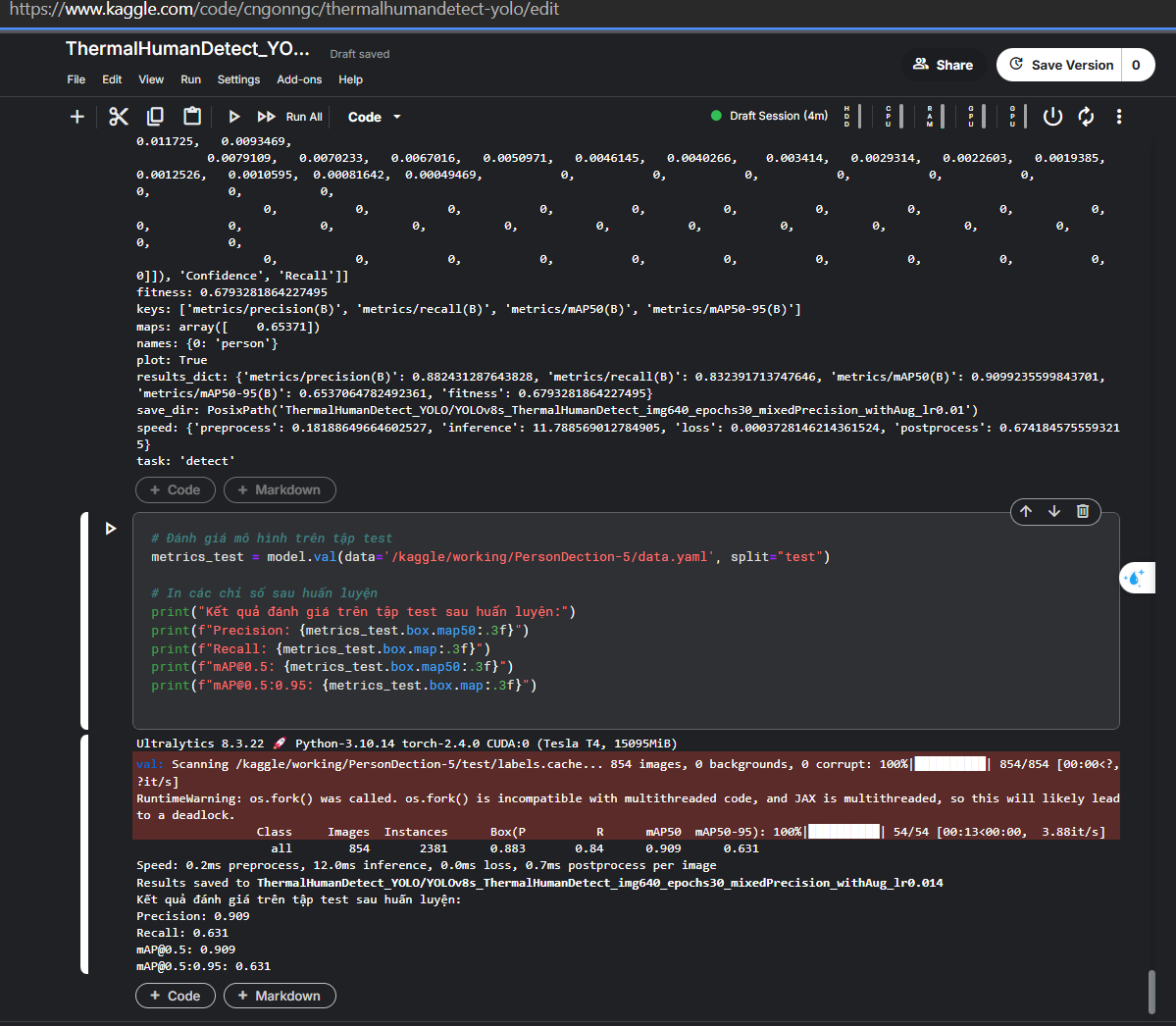
* Yolov8

 **Giai đoạn 1:** Huấn luyện mô hình **Yolov8n (nano)** trong **3 epochs** mà không sử dụng kỹ thuật tăng cường dữ liệu (No Augmentation). Mục tiêu của giai đoạn này là đánh giá khả năng học đặc trưng của mô hình từ dữ liệu ảnh nhiệt ban đầu. Do chỉ chạy trong thời gian ngắn và không có augmentation, giai đoạn này chủ yếu giúp thu thập kết quả cơ bản ban đầu để so sánh với các giai đoạn tiếp theo.

 **Giai đoạn 2:** Mô hình tiếp tục được huấn luyện thêm với **15 epochs** **mà vẫn không sử dụng augmentation.** Giai đoạn này giúp cải thiện đáng kể khả năng phát hiện đối tượng và dự đoán chính xác hơn nhờ việc học sâu hơn từ dữ liệu. Đây là bước trung gian để kiểm tra xem mô hình có thể tự cải thiện mà không cần tăng cường dữ liệu hay không.

 **Giai đoạn 3:** Chuyển sang mô hình **Yolov8m (medium)** và huấn luyện trong **30 epochs** với áp dụng **kỹ thuật** **Augmentation**. Đây là giai đoạn huấn luyện quan trọng nhất, sử dụng augmentation để tăng cường sự đa dạng của dữ liệu, giúp mô hình học tốt hơn và tổng quát hóa hơn cho các tình huống phức tạp. Điều này đặc biệt cần thiết để tránh hiện tượng overfitting và tăng cường khả năng nhận diện đối tượng trong điều kiện thực tế.

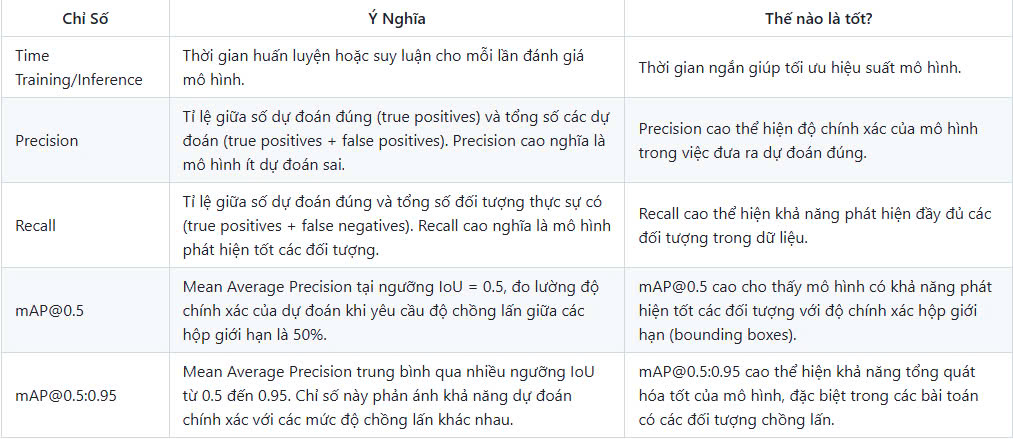




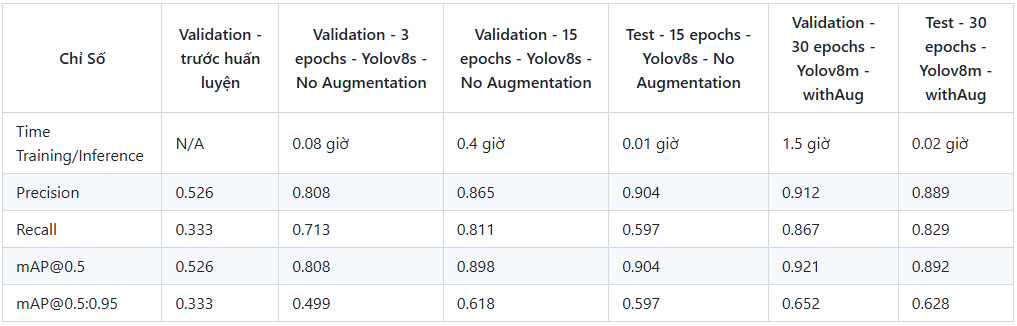
**[Ảnh quá trình training model]**

### **3.1.4 Kết quả của mô hình YOLOv8:**

* **Các metrics (chỉ số sử dụng để đánh giá model)**



**Các metrics (chỉ số sử dụng để đánh giá model)**





**[Bảng kết quả Model sau quá trình Training]**

**Nhận xét kết quả:**

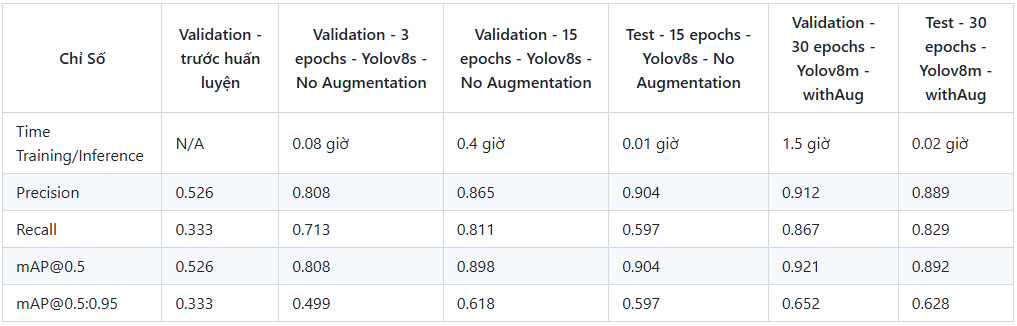
- **Trước khi huấn luyện:** Khi thực hiện đánh giá ban đầu trên tập validation, mô hình YOLOv8 đạt Precision là 0.526, Recall là 0.333, mAP@0.5 là 0.526, và mAP@0.5:0.95 là 0.333. **Đây là các chỉ số từ mô hình YOLOv8 cơ bản chưa được tối ưu hóa.**

- **Sau khi huấn luyện (3 epochs - Yolov8s - No Augmentation) – 0.08h (4.8min):** Sau 3 epochs huấn luyện, các chỉ số cho thấy sự cải thiện đáng kể: Precision tăng lên 0.808, Recall đạt 0.713, mAP@0.5 là 0.808 và mAP@0.5:0.95 là 0.499. Những kết quả này thể hiện rằng mô hình đã học được đặc trưng quan trọng của đối tượng người trong các ảnh nhiệt.

- **Sau khi huấn luyện (15 epochs - Yolov8s - No Augmentation) – 0.4h (24min):** Sau 15 epochs huấn luyện, các chỉ số tiếp tục cải thiện với Precision đạt 0.865, Recall là 0.811, mAP@0.5 đạt 0.898 và mAP@0.5:0.95 đạt 0.618 trên tập validation. Điều này thể hiện rằng mô hình đã học tốt hơn về khả năng phát hiện và nhận diện đối tượng trong các điều kiện phức tạp.

- **Chuyển qua huấn luyện (30 epochs – Yolov8m – with Augmentation) – 1.5h (90min):** Sau khi hoàn tất huấn luyện với mô hình Yolov8m và bổ sung kỹ thuật tăng cường dữ liệu (augmentation), mô hình tiếp tục cải thiện trên tất cả các chỉ số. Precision đạt **0.912**, Recall tăng lên **0.867**, mAP@0.5 đạt **0.921**, và mAP@0.5:0.95 đạt **0.652**. Kỹ thuật tăng cường dữ liệu đã giúp mô hình trở nên mạnh mẽ hơn trong việc nhận diện đối tượng người dưới các điều kiện phức tạp và biến động của dữ liệu.







**Ảnh So sánh 4 chỉ số của 3 models: 3 epochs, 15 epochs, với 30 epochs**

**3.1.5 Đánh giá hiệu năng mô hình trên ảnh nhiệt:**

 **Precision:** Sau huấn luyện, Precision đạt **0.912**, cho thấy mô hình đưa ra các dự đoán chính xác hơn và ít lỗi hơn so với Precision trước huấn luyện (**0.526**).

 **Recall:** Recall cũng cải thiện đáng kể, tăng từ **0.333** lên **0.867**, cho thấy khả năng phát hiện đối tượng của mô hình đã được cải thiện đáng kể.

 **mAP@0.5 và mAP@0.5:0.95:** Sau huấn luyện, mAP@0.5 tăng từ **0.526** lên **0.921**, và mAP@0.5:0.95 tăng từ **0.333** lên **0.652**, cho thấy mô hình đã cải thiện đáng kể khả năng phát hiện chính xác trong các tình huống phức tạp khi đối tượng bị chồng lấn.

**3.1.6 Phân tích kết quả và độ chính xác của mô hình**

- Sau quá trình huấn luyện, mô hình YOLOv8 đã thể hiện sự cải thiện rõ rệt. **Precision** đã tăng từ **0.526** trước huấn luyện lên **0.912** sau khi huấn luyện với 50 epochs và augmentation, cho thấy mô hình dự đoán chính xác hơn và ít sai số hơn. **Recall** cũng tăng từ **0.333** lên **0.867**, minh chứng cho khả năng phát hiện đối tượng tốt hơn, đảm bảo mô hình không bỏ sót các đối tượng quan trọng.

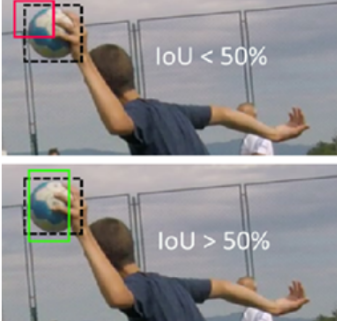
- Các chỉ số **mAP@0.5** và **mAP@0.5:0.95** cũng cho thấy sự cải tiến đáng kể, với **mAP@0.5** tăng từ **0.526** lên **0.921** và **mAP@0.5:0.95** tăng từ **0.333** lên **0.652**. Điều này phản ánh khả năng của mô hình trong việc phát hiện đối tượng chính xác hơn, ngay cả khi đối tượng trong ảnh có mức độ chồng lấn.

- Mặc dù hiệu suất mô hình đã được cải thiện, vẫn có tiềm năng tối ưu hơn nữa, đặc biệt là về tốc độ huấn luyện và khả năng tổng quát hóa trên các dữ liệu phức tạp hơn. Điều này có thể đạt được thông qua việc điều chỉnh thêm các tham số hoặc áp dụng các phương pháp huấn luyện tiên tiến hơn.

**CHƯƠNG 4: KẾT LUẬN VÀ KIẾN NGHỊ**

**4.1 Kết luận :**

Mô hình YOLO (You Only Look Once) đã chứng minh hiệu quả cao trong việc nhận diện người từ ảnh nhiệt. Với khả năng xử lý nhanh và chính xác, YOLO không chỉ giúp xác định vị trí người trong các bối cảnh khác nhau mà còn cải thiện đáng kể độ chính xác so với các phương pháp truyền thống. Việc áp dụng YOLO trong nhận diện từ ảnh nhiệt mang lại nhiều lợi ích, đặc biệt trong các tình huống ánh sáng yếu hoặc môi trường khắc nghiệt, nơi mà các kỹ thuật nhận diện thông thường gặp khó khăn.



Hình 4.1: Biểu diễn cực quan

**4.2 Hướng phát triển :**

Để nâng cao hiệu quả của mô hình YOLO trong tương lai, một số cải tiến và ứng dụng có thể được xem xét:

1. **Tối ưu hóa mô hình**: Thực hiện tinh chỉnh hyperparameters và kiến trúc mạng để tăng cường độ chính xác và giảm thiểu thời gian xử lý.
2. **Mở rộng tập dữ liệu**: Xây dựng và sử dụng các tập dữ liệu đa dạng hơn để cải thiện khả năng tổng quát của mô hình, đặc biệt là trong các điều kiện thời tiết và ánh sáng khác nhau.
3. **Ứng dụng thực tiễn**: Khai thác mô hình YOLO trong các lĩnh vực như an ninh, giám sát và cứu hộ, nơi mà việc phát hiện người nhanh chóng và chính xác là rất quan trọng.
4. **Kết hợp với công nghệ khác**: Tích hợp YOLO với các công nghệ như AI và IoT để phát triển các giải pháp thông minh cho việc giám sát và quản lý an ninh.

Các cải tiến và ứng dụng này sẽ mở ra nhiều cơ hội mới cho mô hình YOLO trong việc nhận diện con người từ ảnh nhiệt và ứng dụng trong thực tiễn.

**TÀI LIỆU THAM KHẢO**

https://www.researchgate.net/publication/333360405\_Human\_Detection\_in\_Thermal\_Imaging\_Using\_YOLO

https://www.superannotate.com/blog/yolo-object-detection

https://github.com/Darijan23/Project\_13

<https://huggingface.co/pitangent-ds/YOLOv8-human-detection-thermal>

https://github.com/ultralytics/yolov5

------Hết------