Shape, square

Description automatically generated**TRƯỜNG ĐẠI HỌC KIẾN TRÚC HÀ NỘI**

**KHOA CÔNG NGHỆ THÔNG TIN**

Logo

Description automatically generated  
**BÁO CÁO BÀI TẬP LỚN**

**CHUYÊN ĐỀ MẠNG MÁY TÍNH VÀ HỆ THỐNG THÔNG TIN**

**ĐỀ TÀI: Phát triển phần mềm phân loại chất lượng hoa quả sử dụng mô hình học sâu YOLO**

**Giảng viên hướng dẫn: TS. Nguyễn Quốc Trung**

**Nhóm sinh viên thực hiện: Nguyễn Văn Vương 2055010297**

**Lớp: 20CN3**

**Hà Nội 2024**

**MỤC LỤC**

[MỞ ĐẦU 4](#_Toc169245205)

[Chương I. TỔNG QUAN 8](#_Toc169245206)

[1.Mô hình YOLO 8](#_Toc169245207)

[1.1 Giới thiệu 8](#_Toc169245208)

[1.2 Lịch sử, cách thức học máy 8](#_Toc169245209)

[1.3 Cách thức học máy của mô hình YOLO 10](#_Toc169245210)

[1.4 Ưu và nhược điểm 12](#_Toc169245211)

[Chương II. TRIỂN KHAI HUẤN LUYỆN MÔ HÌNH NHẬN DIỆN 14](#_Toc169245212)

[1. Lập kế hoạch 14](#_Toc169245213)

[1.1 Giai đoạn 1: Thu thập và chuẩn bị dữ liệu 14](#_Toc169245214)

[1.2 Giai đoạn 2: Tiền xử lý dữ liệu 14](#_Toc169245215)

[1.3 Giai đoạn 3: Huấn luyện mô hình nhận diện vùng chứa hoa quả 14](#_Toc169245216)

[1.4 Giai đoạn 4: Huấn luyện mô hình phân loại hoa quả tươi và hư hại 15](#_Toc169245217)

[1.5 Giai đoạn 5: Kiểm tra và đánh giá mô hình 15](#_Toc169245218)

[1.6 Giai đoạn 6: Triển khai và tinh chỉnh mô hình 15](#_Toc169245219)

[1.7 Giai đoạn 7: Bảo trì và cập nhật mô hình 15](#_Toc169245220)

[2. Thực hiện kế hoạch 16](#_Toc169245221)

[2.1 Thu thập và chuẩn bị dữ liệu 16](#_Toc169245222)

[2.2 Thực hiện huấn luyện mô hình 18](#_Toc169245223)

[2.3 Đánh giá mô hình sau khi huấn luyện 21](#_Toc169245224)

[Chương III: TRIỂN KHAI MÔ HÌNH SỬ DỤNG FLASKAPI 23](#_Toc169245225)

[3.1 Giới thiệu về FLASKAPI 23](#_Toc169245226)

[3.2 Các tính năng của Flask Framework 23](#_Toc169245227)

[3.3 Đánh giá ưu và nhược điểm của Flask 24](#_Toc169245228)

[3.4 Xây dựng chương trình demo: 25](#_Toc169245229)

[3.4.1 Xây dựng một giao diện web đơn giản thực hiện việc upload ảnh 25](#_Toc169245230)

[3.4.2 Xây dựng giao diện đơn giản cho việc cho ra kết quả sau khi model đã dữ đoán: 25](#_Toc169245231)

[3.4.3 Xây dựng scripts sử dụng Flask cho việc triển khai model lên web: 26](#_Toc169245232)

[CHƯƠNG IV: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 30](#_Toc169245233)

[4.1 Kết luận 30](#_Toc169245234)

[4.2 Hướng phát triển 30](#_Toc169245235)

[TÀI LIỆU THAM KHẢO 32](#_Toc169245236)

# MỞ ĐẦU

Ngày nay, với sự phát triển của khoa học công nghệ và các kỹ thuật tiên tiến, hiện đại, việc áp dụng các kỹ thuật đó vào đời sống của con người là thật sự cần thiết, bởi những việc đó không chỉ giúp cho cuộc sống của con người được cải thiện mà còn làm cho xã hội phát triển nhanh chóng, góp phần thúc đẩy sự tiến bộ của khoa học, kỹ thuật cũng như bắt kịp với xu hướng của thời đại.

Một trong những lĩnh vực hiện đại đang dẫn đầu xu hướng hiện nay trong công cuộc phát triển của thời đại 4.0 chính là nghiên cứu về trí tuệ nhân tạo. Trí tuệ nhân tạo là lĩnh vực đang được phát triển nhất trên thế giới, bằng việc sử dụng các phương pháp machine learning hay deep learning,… và một số phương pháp khác đã giúp cho lĩnh vực nghiên cứu về trí tuệ nhân tạo đang dẫn đầu xu hướng không chỉ ở Việt Nam mà còn ra bên ngoài thế giới. Do vậy, trong lĩnh vực nghiên cứu về trí tuệ nhân tạo, tính đến thời điểm hiện tại, đã có nhiều ứng dụng phục vụ đời sống của con người như nhận diện khuôn mặt, phân loại làn đường, trợ lý ảo, nhận dạng chữ viết tay, nhận diện hoa quả tươi và hư hại ,...

Trong bối cảnh phát triển của công nghệ trí tuệ nhân tạo, việc áp dụng mô hình học sâu vào các ứng dụng thực tế ngày càng phổ biến. Báo cáo này trình bày chi tiết việc sử dụng mô hình YOLOv8 để phát triển phần mềm phân loại hoa quả tươi và hư hại. Qua quá trình huấn luyện và đánh giá, mô hình đã chứng minh được hiệu quả cao trong việc nhận diện và phân loại các loại hoa quả, đồng thời cung cấp một giải pháp khả thi cho việc kiểm tra chất lượng nông sản trong ngành công nghiệp thực phẩm.

Chương I. GIỚI THIỆU

1. Bài toán nhận diện vật thể

Một trong những lĩnh vực quan trọng của Trí tuệ nhân tạo (Artificial Intelligence) là thị giác máy (Computer Vision). Computer Vision là một lĩnh vực bao gồm các phương pháp thu nhận, xử lý ảnh kỹ thuật số, phân tích và nhận dạng các hình ảnh, phát hiện các đối tượng, tạo ảnh, siêu phân giải hình ảnh và nhiều hơn vậy. Object Detection có lẽ là khía cạnh sâu sắc nhất của thị giác máy do số lần sử dụng trong thực tế.

CNN được viết tắt của Convolutional Neural Network hay còn được gọi là CNNS mang nơ-ron tích chập, là một trong những mô hình Deep Learning cực kỳ tiên tiến cho phép xây dựng những hệ thống có độ chính xác cao và thông minh. Nhờ khả năng đó, CNN có rất nhiều ứng dụng, đặc biệt là những bài toán cần nhận dạng vật thể (object) trong ảnh. CNN vô cùng quan trọng để tạo nên những hệ thống nhận diện thông minh với độ chính xác cao trong thời đại công nghệ ngày nay.

1. Ứng dụng của bài toán nhận diện vật thể trong đời sống

### 1.1 Giám sát An ninh

* **Phát hiện xâm nhập**: Hệ thống giám sát an ninh sử dụng nhận diện đối tượng để phát hiện các hành vi xâm nhập trái phép vào khu vực được bảo vệ.
* **Nhận diện khuôn mặt**: Công nghệ nhận diện khuôn mặt được tích hợp trong các camera giám sát để nhận diện và theo dõi các cá nhân trong đám đông, hỗ trợ trong việc tìm kiếm tội phạm hoặc người mất tích.

### 1.2 Giao thông và An toàn Đường bộ

* **Phát hiện vi phạm giao thông**: Camera giao thông có thể nhận diện các hành vi vi phạm như vượt đèn đỏ, chạy quá tốc độ, và không đội mũ bảo hiểm.
* **Quản lý đỗ xe**: Hệ thống nhận diện biển số xe giúp quản lý bãi đỗ xe tự động, từ việc ghi nhận xe vào ra cho đến việc phát hiện các vi phạm về đỗ xe.

## 2. Y tế

* **Phát hiện bệnh lý**: Nhận diện đối tượng trong hình ảnh y tế như X-quang, MRI, và CT scan giúp bác sĩ phát hiện sớm các bệnh lý như ung thư, tổn thương não, và các vấn đề về tim mạch.
* **Theo dõi bệnh nhân**: Camera giám sát trong bệnh viện có thể theo dõi hành vi của bệnh nhân, phát hiện các tình huống nguy hiểm như ngã hoặc co giật.

## 3. Công nghiệp và Sản xuất

* **Kiểm tra sản phẩm**: Hệ thống nhận diện đối tượng giúp tự động kiểm tra các sản phẩm trong dây chuyền sản xuất để phát hiện lỗi và loại bỏ các sản phẩm không đạt tiêu chuẩn.
* **Phân loại sản phẩm**: Trong ngành công nghiệp thực phẩm, nhận diện đối tượng giúp phân loại hoa quả, rau củ, và các sản phẩm khác dựa trên hình dạng, kích thước, và màu sắc.

### 3.2 Quản lý Kho Hàng

* **Theo dõi hàng hóa**: Sử dụng công nghệ nhận diện đối tượng để theo dõi và quản lý hàng hóa trong kho, từ việc kiểm tra tồn kho cho đến việc giám sát quá trình vận chuyển.

## 4. Bán lẻ

* **Cửa hàng không thu ngân**: Các cửa hàng như Amazon Go sử dụng công nghệ nhận diện đối tượng để theo dõi các sản phẩm mà khách hàng lấy và tính tiền tự động khi họ rời cửa hàng.
* **Phân tích hành vi khách hàng**: Camera giám sát trong cửa hàng có thể phân tích hành vi mua sắm của khách hàng để tối ưu hóa bố cục cửa hàng và chiến lược tiếp thị.
* **Kiểm tra kệ hàng**: Hệ thống nhận diện đối tượng có thể theo dõi tình trạng kệ hàng, phát hiện các sản phẩm sắp hết và thông báo cho nhân viên kịp thời bổ sung.

## 5. Giao thông và Vận tải

* **Hệ thống ADAS**: Hệ thống hỗ trợ lái xe tiên tiến (Advanced Driver Assistance Systems) sử dụng nhận diện đối tượng để phát hiện các vật cản, người đi bộ, và phương tiện khác, hỗ trợ người lái trong việc tránh va chạm.
* **Xe tự hành**: Các phương tiện tự hành sử dụng công nghệ nhận diện đối tượng để điều hướng và xử lý các tình huống giao thông phức tạp.
* **Điều khiển đèn giao thông thông minh**: Nhận diện đối tượng giúp tối ưu hóa thời gian thay đổi đèn tín hiệu dựa trên lưu lượng xe cộ, giảm tắc nghẽn giao thông.

## 6. Thể thao và Giải trí

* **Phân tích trận đấu**: Công nghệ nhận diện đối tượng được sử dụng để theo dõi và phân tích chuyển động của cầu thủ trong các trận đấu thể thao, cung cấp dữ liệu chi tiết cho huấn luyện viên và bình luận viên.
* **Phát lại tức thời**: Hệ thống nhận diện đối tượng giúp tạo ra các đoạn phát lại tức thời (instant replay) trong các sự kiện thể thao, tập trung vào các tình huống quan trọng.
* **Trải nghiệm tương tác**: Nhận diện đối tượng giúp tăng cường trải nghiệm thực tế ảo (VR) và thực tế tăng cường (AR) bằng cách nhận diện và tương tác với các đối tượng trong môi trường thực tế.

## 7. Nông nghiệp

* **Phát hiện bệnh cây trồng**: Hệ thống nhận diện đối tượng giúp phát hiện sớm các dấu hiệu bệnh lý trên cây trồng, hỗ trợ nông dân trong việc quản lý và bảo vệ mùa màng.
* **Theo dõi vật nuôi**: Camera giám sát có thể theo dõi hành vi và tình trạng sức khỏe của vật nuôi, phát hiện các dấu hiệu bất thường.
* **Máy thu hoạch tự động**: Sử dụng nhận diện đối tượng để xác định và thu hoạch các loại hoa quả và rau củ một cách tự động, tăng hiệu suất và giảm chi phí lao động.

**Mục tiêu đề tài**

Xây dựng ứng dụng phân loại hoa quả tươi và hư hại qua hình ảnh thu được từ camera nhằm đáp ứng nhu cầu quản lý chất lượng sản phẩm tại các nhà máy chế biến hoa quả nói chung và quản lý chất lượng hoa quả tại các cửa hàng bán lẻ nói riêng. Trong quá trình đó, để tăng tính chính xác và hiệu quả, chúng tôi nghiên cứu thêm phương pháp phân loại hoa quả để tích hợp hệ thống nhằm phân biệt rõ ràng giữa hoa quả tươi và hoa quả hư hại.

* Tìm hiểu và sử dụng mô hình YOLO trong việc huấn luyện mô hình phân loại hoa quả.
* Tìm hiểu các thuật toán học máy khác như CNN và các công cụ hỗ trợ xử lý ảnh.
* Xây dựng một mô hình học máy cho phép nhận dạng và phân loại hoa quả thông qua các camera được thiết lập tại cơ sở chế biến hoặc cửa hàng.
* Áp dụng mô hình phân loại để xác định tình trạng tươi hay hư hại của hoa quả.
* Áp dụng các framework UI, Mobile để xây dựng giao diện ứng dụng thân thiện với người dùng.
* Vận dụng các kỹ thuật để xây dựng ứng dụng web và mobile cho việc quản lý và theo dõi chất lượng hoa quả.
* Đánh giá kết quả và đề xuất phương hướng phát triển để nâng cao độ chính xác và hiệu quả của hệ thống.

**Phương pháp nghiên cứu**

Ứng dụng các mô hình học máy kết hợp với một số phương pháp khác để xây dựng được ứng dụng có thể áp dụng vào thực tiễn, đặc biệt là có thể áp dụng trong việc quản lý chất lượng hoa quả tại các cửa hàng bán lẻ nói riêng

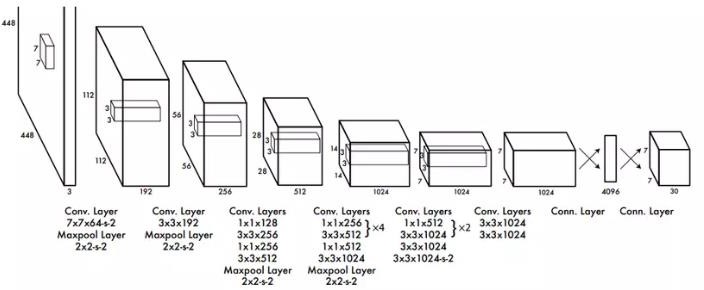
**II. Cơ sở lý thuyết**

## 2.1.Mô hình YOLO

## 1.1 Giới thiệu

YOLO, 1 từ viết tắt cho ‘You only look once’ (Bạn chỉ nhìn một lần) là một thuật toán phân loại đối tượng chia hình ảnh thành một hệ thống lưới. Mỗi ô trong lưới chịu trách nhiệm phân loại các đối tượng trong chính nó. YOLO là một trong những thuật toán phân loại đối tượng nổi tiếng nhất do tốc độ và độ chính xác của nó.

YOLO là một mô hình mạng CNN cho việc phân loại, nhận dạng, phân loại đối tượng. YOLO được tạo ra từ việc kết hợp giữa các convolutional layers và connected layers. Trong đó các convolutional layers sẽ trích xuất ra các feature của ảnh, còn full- connected layers sẽ dự đoán ra xác suất đó và tọa độ của đối tượng.



Hình 1.1: Mô hình mạng CNN được sử dụng trong YOLO

### 1.2 Lịch sử, cách thức học máy

Lịch sử của YOLO

YOLOv1 đã được Joseph Redmon phát hành dưới dạng nghiên cứu. Tờ báo có tiêu đề là ‘You Only Look Once: Unified, Real-Time Object Detection’ được phát hành vào ngày 8 tháng 6 năm 2015.

YOLOv2 đã được phát hành bởi Joseph Redmon - tác giả ban đầu của YOLO và Ali Farhadi, cùng nhau xuất bản ‘YOLO9000: Better, Faster, Stronger’ vào ngày 25 tháng 12 năm 2016.

YOLOv3 được cải tiến dựa trên YOLOv2 bởi 2 tác giả ban đầu YOLOv2, họ cùng nhau xuất bản ‘YOLOv3: An incremental Improvement’ vào ngày 8 tháng 4 năm 2018.

YOLOv4 được phát triển bởi Aley Bochoknovskiy, Chien-Yao Wang, and Hong- Yuan Mark Liao vì tác giả ban đầu của YOLO đã bế tắc khi phát hành phiên bản mới,

tờ báo có tiêu đề ‘YOLOv4: Optimal Speed and Accuracy of Object Detection’ vào ngày 23 tháng 4 năm 2020.

YOLOv5 được phát hành ngay sau khi YOLOv4 được phát triển dựa vào framework Pytorch bởi tác giả Glenn Jocher, mã nguồn YOLOv5 có sẵn trên GitHub được phát hành ngày 18 tháng 5 năm 2020.

Được sử dụng phổ biến gần đây nhất là YOLOv8 có nhiều tham số hơn so với các phiên bản tiền nhiệm như YOLOv5, nhưng ít tham số hơn so với YOLOv6. Nó cung cấp khoảng 33% mAP nhiều hơn cho các mô hình kích thước n và mAP lớn hơn nói chung.



Hình 1.2 : Kích thước các mô hình YOLOv8

1.3 Cách thức học máy của mô hình YOLO

Hình ảnh đầu vào sẽ được Thuật toán YOLO chia thành SxS thường thì sẽ 3x3, 7x7, 9x9,...

Với dữ liệu đầu vào là 1 ảnh, đầu ra mô hình sẽ là một ma trận 3 chiều có kích thước SxSx(5xB+C) với số lượng tham số mỗi ô là (5xB+C) với B là số lượng Bounding Box, C là lớp mà mỗi ô cần dự đoán.

Ta có bounding box gồm 5 thành phần dự đoán (x,y,w,h, confidence) với (x,y) là tọa độ tâm bouding box so với giới bound của ô lưới, (w,h) là chiều rộng, chiều cao được dự đoán so với toàn bộ hình ảnh,và cuối cùng dự đoán confidence đại diện cho IOU (Intersection over Union (tạm dịch: hợp cả 2) giữa vị trí được dự đoán và vị trí cố định cho trước (predicted box, ground truth box).

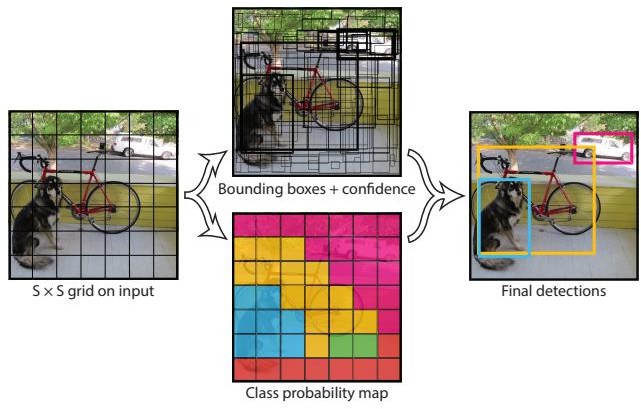
Từng ô lưới dự đoán xác suất lớp (class): Pr(Classi|Object)

Confidence score: nếu không có đối tượng trong ô, thì điểm sẽ là 0, còn không thì bằng Pr(Object)\*IOU



Hình 1.3: Công thức tính độ chính xác phân loại ra đối tượng

Theo như công thức trên ta sẽ ra xác suất final detection sẽ dự đoán và bao các đối tượng trong hình ảnh đó như hình minh họa bên dưới:



Hình 1.4: Phân loại nhận diện đối tượng theo công thức

### Ưu và nhược điểm

* Ưu điểm:

Tốc độ cao:

* YOLO được thiết kế để thực hiện việc phân loại đối tượng trong thời gian thực với tốc độ rất nhanh. Điều này là do toàn bộ quá trình phân loại đối tượng được thực hiện trong một lần duy nhất (one pass) qua mạng neural, không giống như các phương pháp khác đòi hỏi nhiều bước xử lý.

Hiệu quả tính toán:

* YOLO tối ưu hóa về mặt tài nguyên và có thể chạy trên các phần cứng giới hạn như GPU di động hoặc các thiết bị nhúng, điều này làm cho nó trở thành một lựa chọn lý tưởng cho các ứng dụng yêu cầu tính toán thời gian thực.

Độ chính xác cao với các đối tượng lớn:

* YOLO thường hoạt động rất tốt với các đối tượng lớn và có kích thước rõ ràng trong ảnh. Điều này giúp nó đạt được độ chính xác cao trong nhiều trường hợp.

Toàn diện và nhất quán:

* YOLO nhìn toàn bộ hình ảnh trong một lần, giúp cho việc phân loại đối tượng được thực hiện một cách toàn diện và nhất quán hơn so với các phương pháp chia nhỏ hình ảnh ra từng phần.
* Nhược điểm

Độ chính xác thấp với các đối tượng nhỏ hoặc chồng chéo:

* YOLO có thể gặp khó khăn trong việc phân loại các đối tượng nhỏ hoặc các đối tượng bị chồng chéo nhau. Điều này do kiến trúc của YOLO sử dụng các lưới dự đoán lớn, khó nhận diện được các chi tiết nhỏ.

Độ chính xác thấp hơn với các đối tượng có tỷ lệ khung hình đặc biệt:

* Khi các đối tượng có hình dạng hoặc tỷ lệ khung hình không bình thường, YOLO có thể gặp khó khăn trong việc xác định đúng khung bao (bounding box).

Khả năng tổng quát hóa hạn chế:

* YOLO có thể gặp khó khăn trong việc tổng quát hóa các đối tượng trong môi trường khác biệt với dữ liệu huấn luyện. Điều này có nghĩa là khi áp dụng trong các môi trường thực tế với điều kiện ánh sáng, góc nhìn hoặc cảnh quan khác biệt, hiệu suất có thể giảm sút.

Không phù hợp với các ứng dụng yêu cầu độ chính xác cao nhất:

* Mặc dù tốc độ cao, nhưng YOLO không phải luôn đạt được độ chính xác cao nhất so với các phương pháp phân loại đối tượng khác như Faster R-CNN hoặc Mask R-CNN, đặc biệt trong các bài toán yêu cầu độ chính xác rất cao.

## 

# Chương II. TRIỂN KHAI HUẤN LUYỆN MÔ HÌNH NHẬN DIỆN

* Để có thể tiến hành huấn luyện mô hình, ta có thể chia giai đoạn nhận diện hoa quả tươi và hư hại thành các giai đoạn chính, đó là:

1. **Lập kế hoạch**

**1.1 Giai đoạn 1: Thu thập và chuẩn bị dữ liệu**

* + Thu thập hình ảnh của hoa quả tươi và hoa quả hư hại từ nhiều nguồn khác nhau.
  + Gắn nhãn (label) cho các hình ảnh thu thập được để xác định tình trạng tươi hoặc hư hại của từng mẫu hoa quả.
  + Chia tập dữ liệu thành các phần huấn luyện, kiểm tra và đánh giá.

**1.2 Giai đoạn 2: Tiền xử lý dữ liệu**

* + Tiến hành các bước tiền xử lý hình ảnh như cắt, thay đổi kích thước, và tăng cường dữ liệu (augmentation) để tăng tính đa dạng của tập dữ liệu huấn luyện.
  + Chuẩn hóa dữ liệu để đảm bảo tính đồng nhất trong quá trình huấn luyện mô hình.

**1.3 Giai đoạn 3: Huấn luyện mô hình nhận diện vùng chứa hoa quả**

* + Sử dụng mô hình YOLOv8 để huấn luyện nhận diện vùng chứa hoa quả trong hình ảnh.
  + Điều chỉnh siêu tham số (hyperparameters) của mô hình để đạt hiệu suất tối ưu.

**1.4 Giai đoạn 4: Huấn luyện mô hình phân loại hoa quả tươi và hư hại**

* + Sử dụng các thuật toán học sâu (deep learning) như CNN để huấn luyện mô hình phân loại tình trạng của hoa quả.
  + Kết hợp các đặc trưng trích xuất từ hình ảnh vùng chứa hoa quả để phân loại chúng thành tươi hoặc hư hại.

**1.5 Giai đoạn 5: Kiểm tra và đánh giá mô hình**

* + Sử dụng tập kiểm tra để đánh giá độ chính xác và hiệu suất của mô hình đã huấn luyện.
  + Tính toán các chỉ số đánh giá như độ chính xác (accuracy), độ nhạy (recall), độ đặc hiệu (specificity), và F1-score để xác định hiệu suất của mô hình.

**1.6 Giai đoạn 6: Triển khai và tinh chỉnh mô hình**

* + Triển khai mô hình vào hệ thống thực tế và theo dõi hiệu suất trong môi trường thực tế.
  + Thu thập phản hồi và dữ liệu mới từ hệ thống để tiếp tục tinh chỉnh và cải tiến mô hình.
  1. **Giai đoạn 7: Bảo trì và cập nhật mô hình**
  + Liên tục cập nhật tập dữ liệu với các mẫu mới để đảm bảo mô hình không bị lỗi thời.
  + Huấn luyện lại mô hình định kỳ và triển khai các phiên bản cập nhật để duy trì hiệu suất cao.

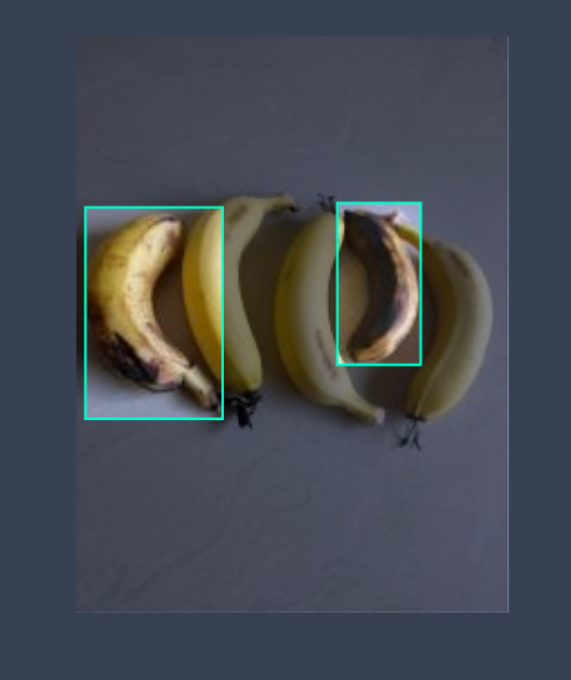
Thông qua các giai đoạn này, mô hình phân loại hoa quả tươi và hư hại sẽ được huấn luyện và tối ưu hóa để đạt được độ chính xác và hiệu suất cao nhất trong việc phân loại tình trạng của hoa quả.

1. **Thực hiện kế hoạch**

### 2.1 Thu thập và chuẩn bị dữ liệu

Trong báo cáo này, tiến hành lấy dữ liệu của 4400 hình ảnh chứa hoa quả tươi và hư hại , trong đó 70% hình ảnh dùng để training, 20% hình ảnh để làm valid và 10% hình ảnh để test. Với mỗi hình ảnh đã được đánh dấu hoa quả tươi và hư hại sẽ tạo ra một tập tin .txt là nhãn của các hoa quả tươi và hư hại tương ứng. Trong tập tin txt này chứa thông tin để nhận dạng hoa quả tươi và hư hại của hình ảnh đó bao gồm lớp, tọa độ x, y, chiều dài và chiều cao.

* Các dữ liệu được tìm kiếm và thu thập từ các trang web như Kaggle, Roboflow, kết hợp với các hình ảnh trên Google…



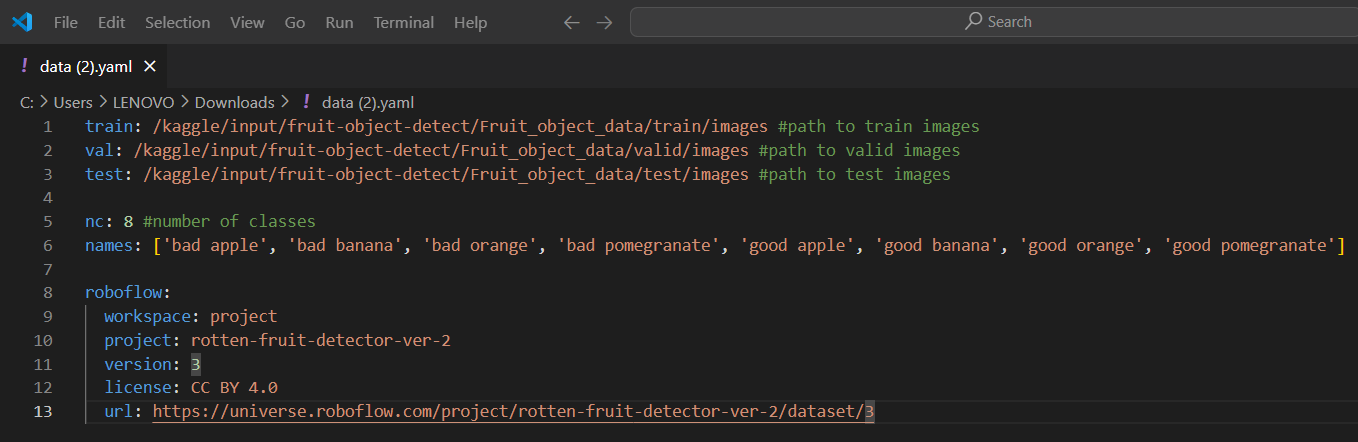
Hình 2.1 : Gán nhãn cho đối tượng

**2.2 Thực hiện huấn luyện mô hình**

Sau khi đã thu thập đủ dữ liệu nhãn cho các hình ảnh. Cuối cùng, sẽ thực hiện huấn luyện dữ liệu cho việc học máy để đảm bảo rằng với mỗi dữ liệu đầu vào là hình ảnh bất kỳ có chứa hoa quả tươi và hư hại đều có thể nhận dạng được. Cụ thể việc huấn luyện dữ liệu được thực hiện như sau:

Tất cả quá trình học máy được thực hiện trên google colab. Đầu tiên, tiến hành tải mã nguồn của YOLOv8 trên github. Đây là mã nguồn mở nên có thể sử dụng cho việc huấn luyện dữ liệu cho phần học máy với các tập dữ liệu đã được thu thập từ trước cùng với một số chỉnh sửa để phù hợp với quá trình nghiên cứu cũng như trong việc nhận dạng vùng có hoa quả tươi và hư hại .

Chỉnh sửa tập tin “data.yaml” cùng với số lớp được khai báo là 8, ở trong báo cáo này chúng ta sẽ nhận diện 4 loại hoa quả phổ biến đó là: quả táo, quả chuối, quả cam và quả lựu với mỗi loại được chia ra 2 class 1 class quả tươi và một class quả hư hại , cấu hình đường dẫn đến thư mục chứa tập tin dữ liệu để bắt đầu huấn luyện. Đó là công việc cần thiết để có thể huấn luyện dữ liệu dựa vào tập dữ liệu đã thu thập trước đó.

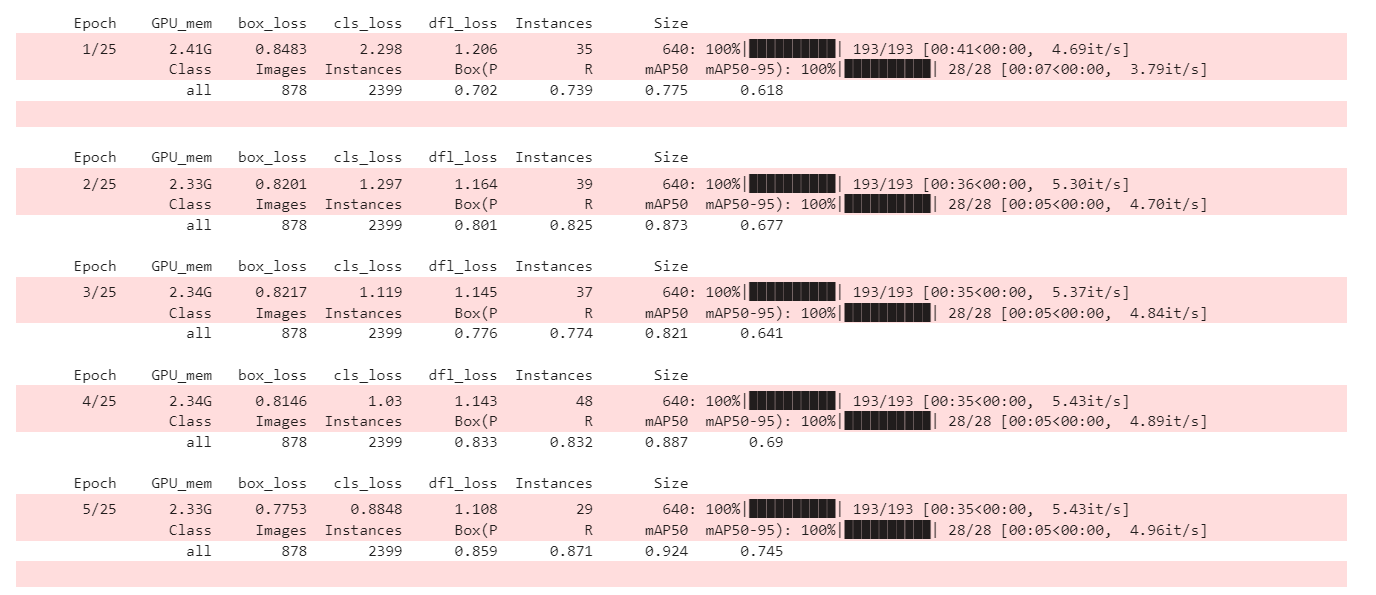


Hình 2.2: Nội dung tập tin “data.yaml”

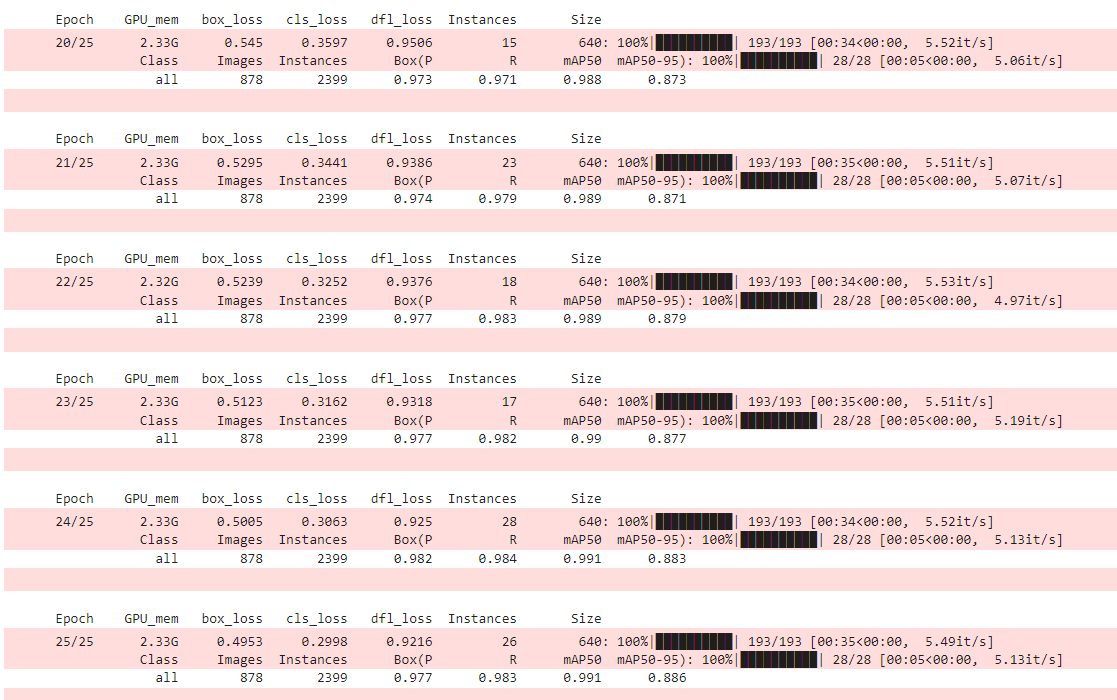
Sau khi chỉnh sửa các nội dung cần thiết trong tập tin, bắt đầu cho quá trình huấn luyện để nhận dạng hoa quả tươi và hư hại được diễn ra, để quá trình học máy được thực hiện cho ra kết quả với độ chính xác cao, thiết lập một số thông số cho quá trình học máy như sau: vì đây là quá trình học máy dựa vào tập dữ liệu gồm có hình ảnh và nhãn của từng hình ảnh tương ứng nên thiết lập thông số của “img” là 640 (do tất cả các hình ảnh trong tập dữ liệu có kích thước không lớn hơn 640px), “batch” là 16 (trong quá trình huấn luyện, số lượng ảnh huấn luyện được cùng một lúc là 16), “epochs” là 25 (số lần lặp lại quá trình huấn luyện là 25), với các thông số được thiết lập ở trên đảm bảo cho quá trình huấn luyện dữ liệu diễn ra cho kết quả với độ chính xác là cao nhất, quá trình học máy được thể hiện như hình:



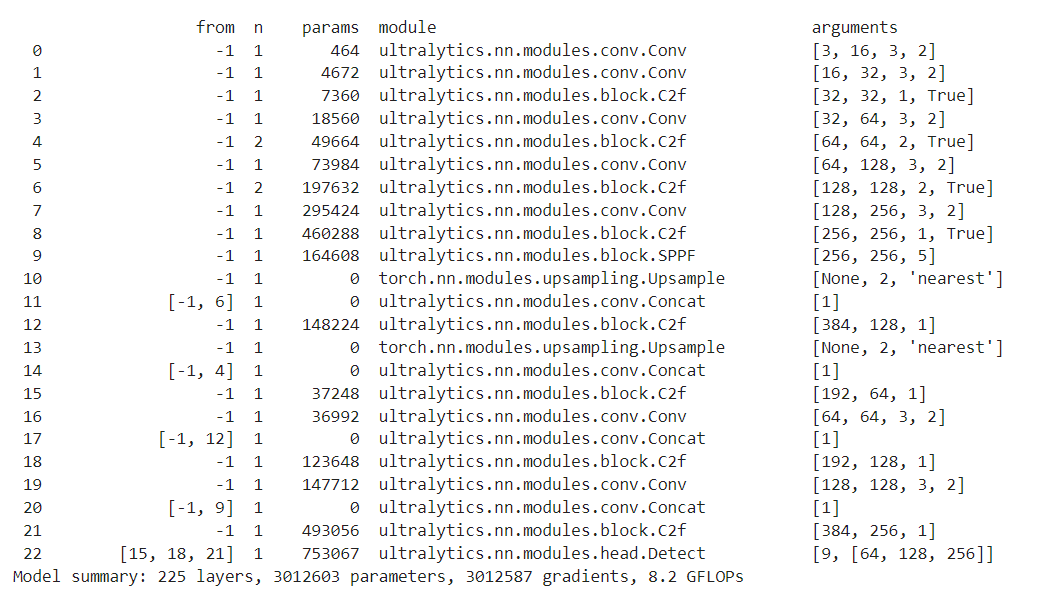
Hình 2.3: Các thông số được sử dụng để huấn luyện



Hình 2.4 : Quá trình huấn luyện trong 5 epochs đầu



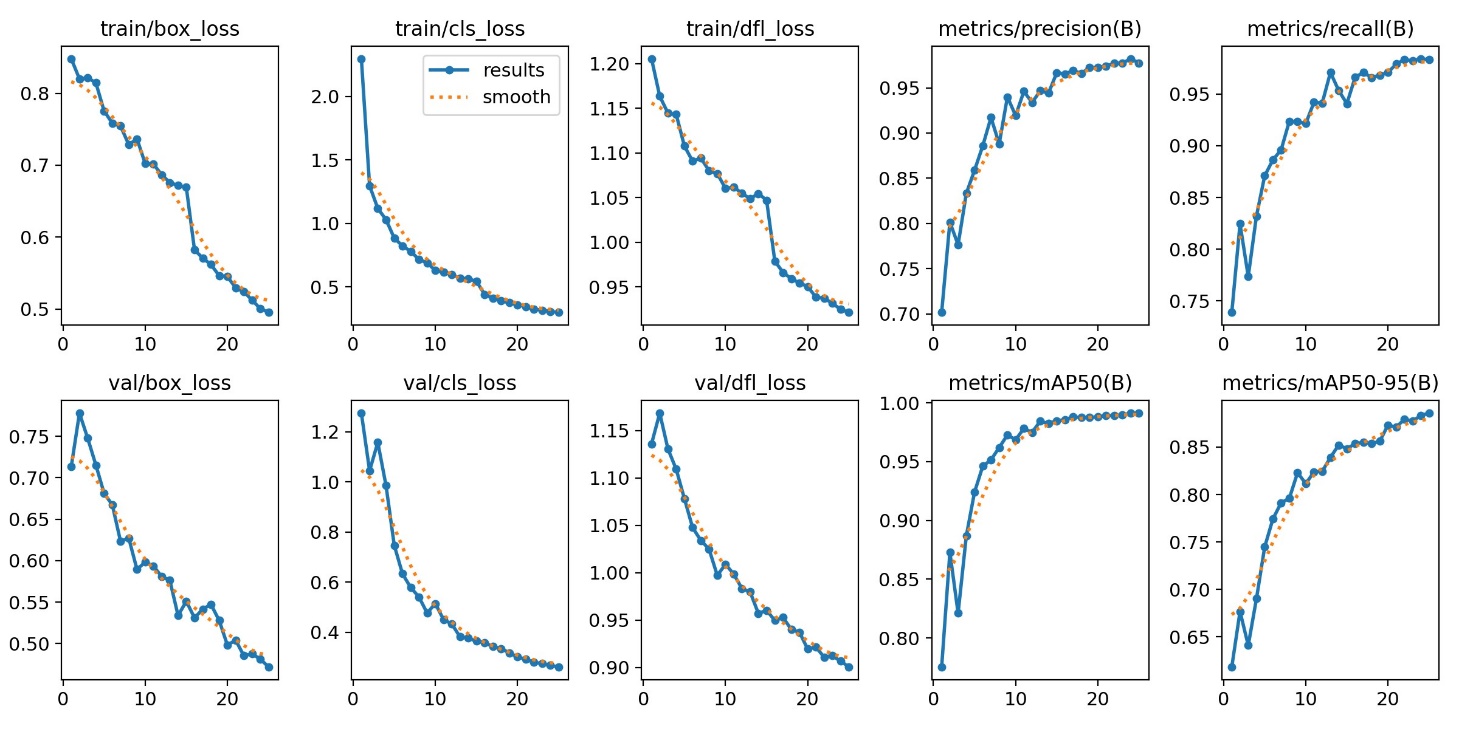
Hình 2.5 : Quá trình huấn luyện trong 5 epochs cuối



Hình 2.6 : Kiến trúc của mô hình với các layer và parameters

* 1. **Đánh giá mô hình sau khi huấn luyện**

Sau khi huấn luyện học máy với việc xác định hoa quả tươi và hư hại , kết quả được thể hiện như sau:

 Hình 2.7 : Biểu đồ đánh giá kết quả huấn luyện dữ liệu

Theo hình trên, nhận thấy rằng mức độ mất mát dữ liệu giảm dần tiệm cận về 0, nghĩa là trong quá trình lặp huấn luyện dữ liệu thì giá trị mất mát giảm dần do trong

quá trình huấn luyện đã dần nhận dạng được trái cây thông qua các lần lặp lại trong quá trình huấn luyện dữ liệu, tương tự với giá trị của đối tượng trong biểu đồ mất mát dữ liệu.

# Chương III: TRIỂN KHAI MÔ HÌNH SỬ DỤNG FLASKAPI

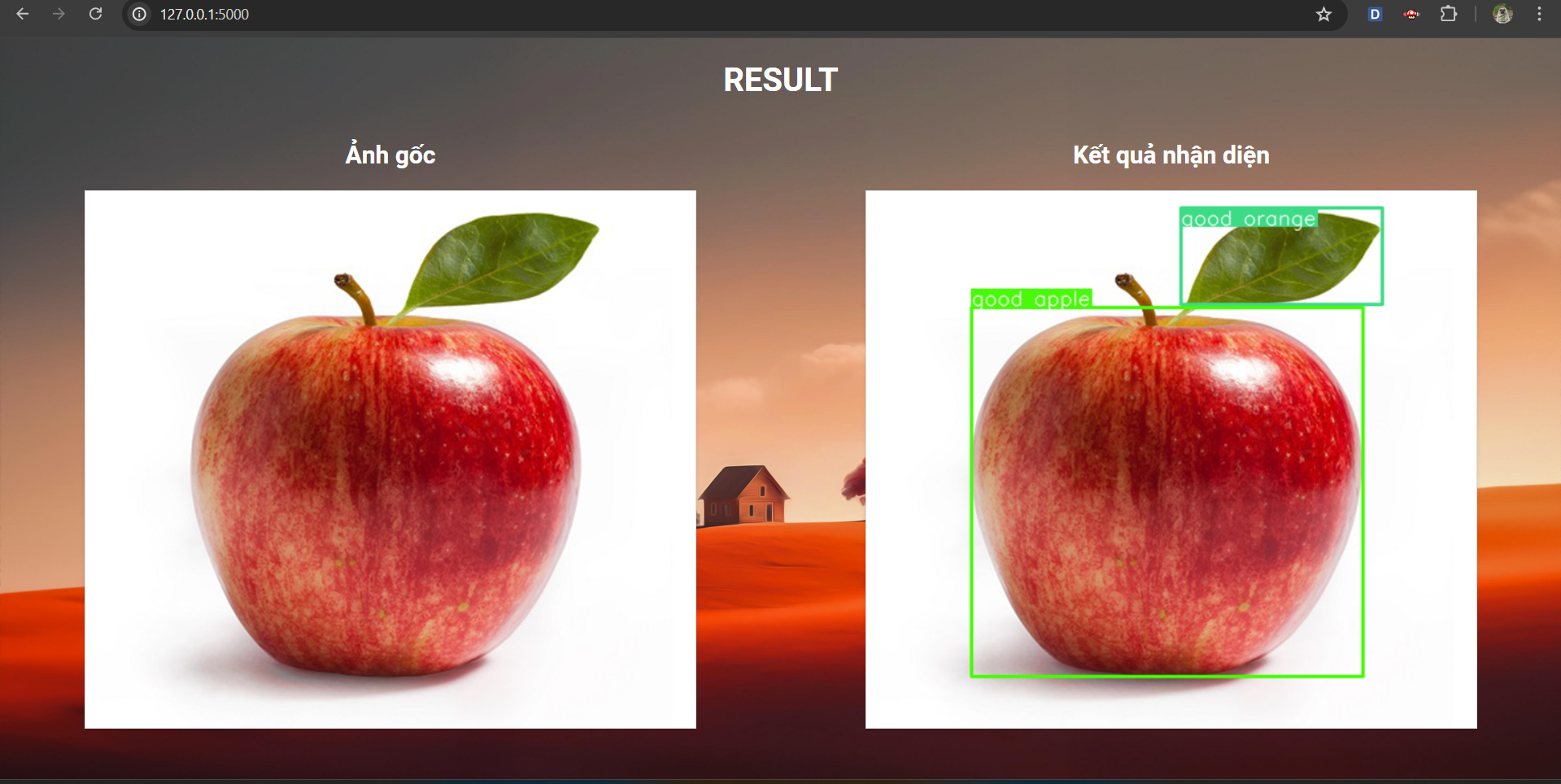
* 1. **Giới thiệu về FLASKAPI**
* Flask là loại framework web phổ biến được viết bằng trình lập ngôn ngữ Python. Công nghệ thường được sử dụng để xây dựng trang web từ những ứng dụng đơn giản đến những hệ thống phức tạp hơn.
* Flask được thiết kế để hoạt động và mở rộng một cách, đồng thời nó cũng cung cấp các công cụ và thư viện cần thiết để phát triển ứng dụng web hiệu quả. Flask cũng có cộng đồng sáng tạo và hỗ trợ mạnh mẽ từ cộng đồng Python.
  1. **Các tính năng của Flask Framework**

Flask Framework sở hữu một số tính năng quan trọng mà nhà phát triển thường sử dụng để xây dựng hiệu ứng web. Dưới đây là một số tính năng chính của Flask:

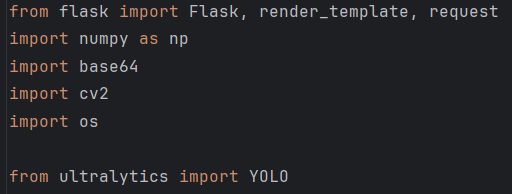
* Nhẹ và dễ sử dụng: Công nghệ có cấu trúc nhẹ nhàng và mã nguồn dễ đọc, giúp người phát triển dễ dàng tiếp cận và tùy chỉnh theo nhu cầu cụ thể của họ.
* Định tuyến linh hoạt: Flask cung cấp cơ chế hoạt động định tuyến, cho phép người phát triển xác định các mẫu URL và phân bổ chúng cho các hàm xử lý tương ứng. Điều này giúp quản lý và xử lý yêu cầu HTTP một cách hiệu quả.
* Công cụ mẫu: Flask tích hợp Jinja2, đây là một loại trình biên dịch mẫu mạnh mẽ cho phép tạo ra các giao diện người dùng.
* Được mở rộng rộng rãi: Mặc dù mang đặc điểm rút gọn nhưng Flask vẫn có khả năng mở rộng mạnh mẽ thông qua việc sử dụng các tiện ích và thư viện của cộng đồng. Người dùng có thể phân tích các tính năng như xác thực, đăng nhập, điều hướng, cơ sở dữ liệu tương tác và nhiều tính năng khác.
* Máy chủ phát triển tích hợp: Flask cung cấp máy chủ phát triển hợp đồng, giúp người phát triển dễ dàng kiểm tra và phát triển ứng dụng mà không cần cấu hình bổ sung.
* Gửi yêu cầu RESTful: Flask hỗ trợ xây dựng API và các ứng dụng RESTful theo cách hoạt động và hiệu quả.
* Cộng đồng lớn và tích cực: Số lượng người dùng Flask rất đông và luôn nhận được hỗ trợ mạnh mẽ từ cộng đồng Python, điều này giúp người phát triển tìm kiếm thông tin và tài liệu một cách dễ dàng.
  1. Đánh giá ưu và nhược điểm của Flask
* Ưu điểm
* Linh hoạt: Flask được thiết kế để linh hoạt và dễ dàng tùy chỉnh theo nhu cầu cụ thể của từng dự án.
* Dễ học và sử dụng: Flask có mã cấu trúc đơn giản, dễ hiểu và tài liệu hướng dẫn phong phú, giúp người dùng dễ dàng tiếp cận và sử dụng.
* Đa chức năng: Flask cho phép tích hợp nhiều công nghệ và thư viện bổ sung khác nhau, từ các công cụ front-end cho cơ sở dữ liệu và mạng dịch vụ.
* Mạnh mẽ cho ứng dụng nhỏ đến trung bình: Flask thích hợp để xây dựng các trang web ứng dụng với quy mô từ nhỏ đến trung bình với hiệu suất tốt.
* Nhược điểm
* Khả năng mở rộng: Flask không cung cấp sẵn các tính năng lớn hơn như các khung, điều này có thể làm tăng mức độ phức tạp khi cần tích hợp nhiều tính năng mở rộng.
* Không phù hợp cho ứng dụng lớn: Bình không phù hợp cho các ứng dụng lớn với quy mô phức tạp và yêu cầu hiệu suất cao.
  1. **Xây dựng chương trình demo:**
     1. Xây dựng một giao diện web đơn giản thực hiện việc upload ảnh



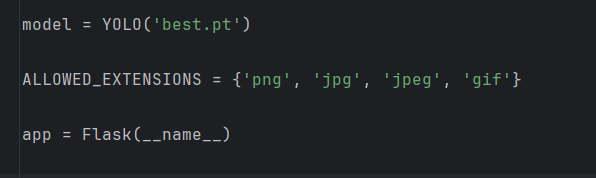
* + 1. Xây dựng giao diện đơn giản cho việc cho ra kết quả sau khi model đã dữ đoán:



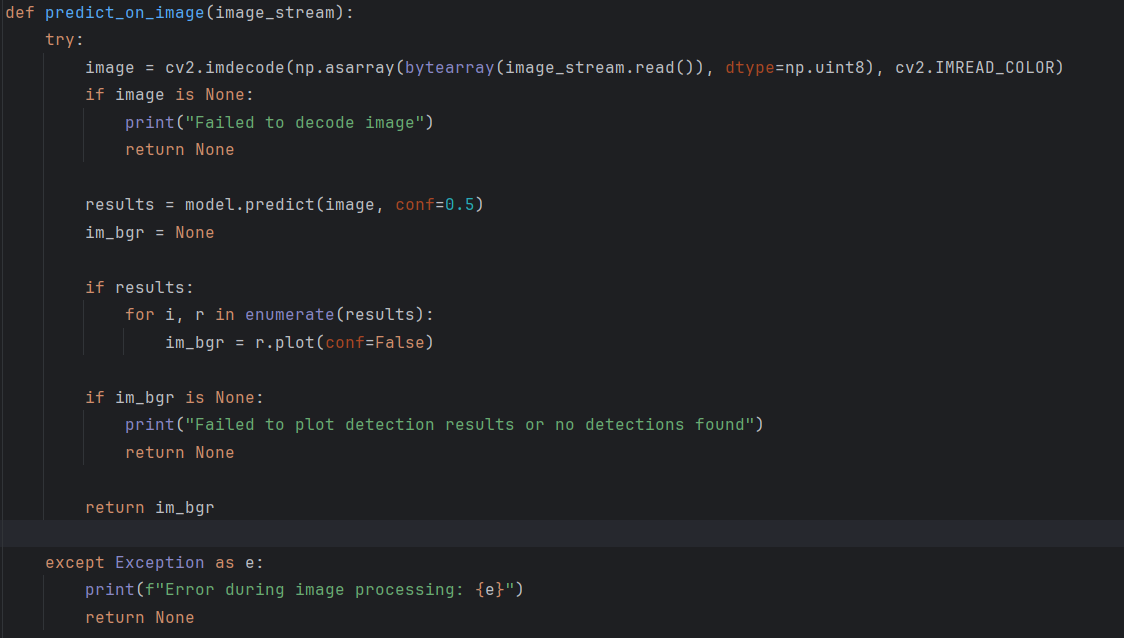
* + 1. Xây dựng scripts sử dụng Flask cho việc triển khai model lên web:
       1. Import các thư viện cần thiết:



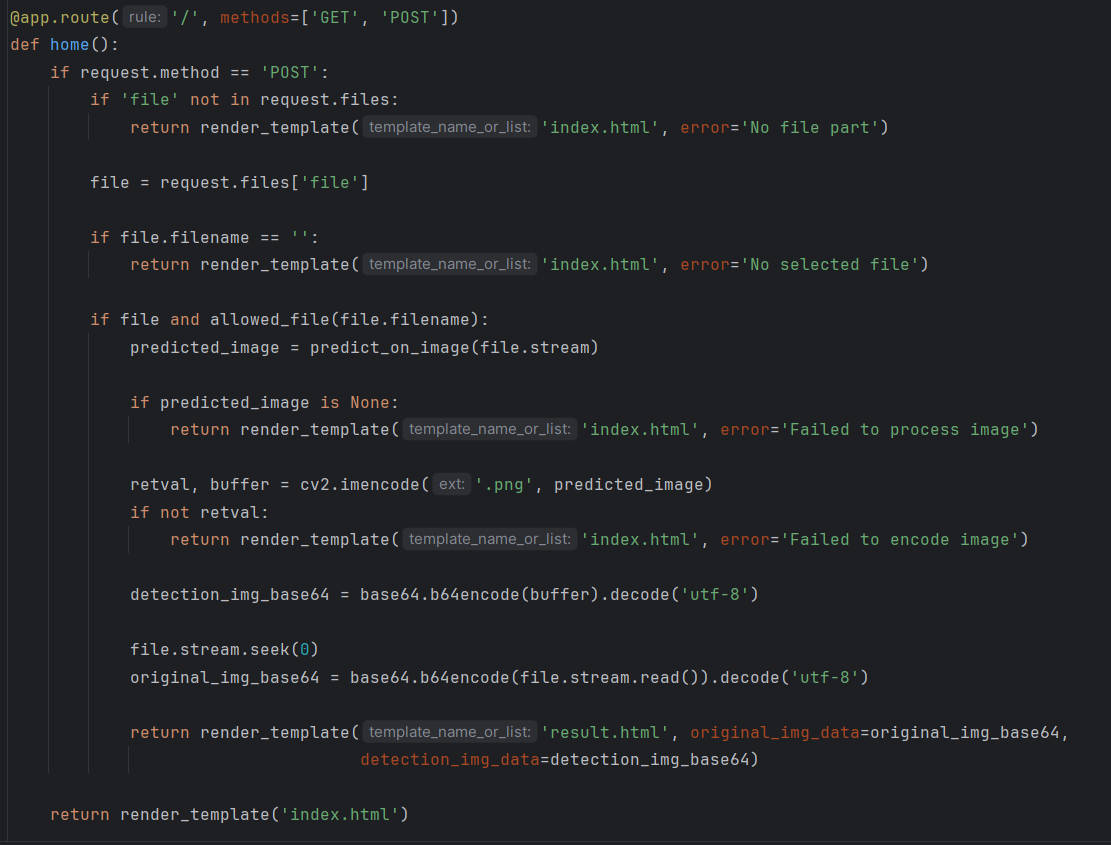
* + - 1. Khởi tạo model từ đã được huấn luyện và Flask:



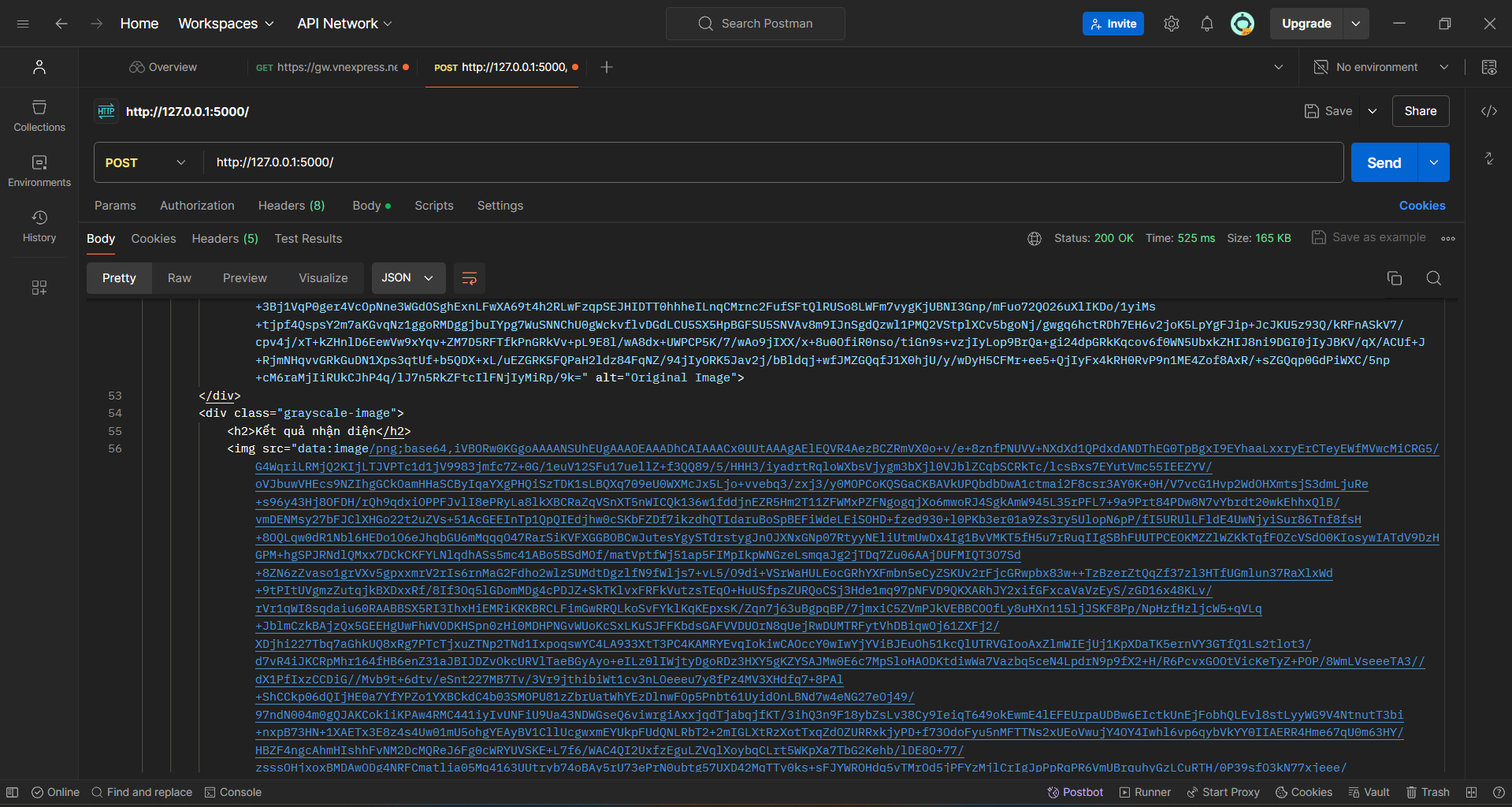
* + - 1. Chạy model dự đoán trên hình ảnh và trả về hình ảnh đã được dự đoán:



* + - 1. Xây dựng hai phương thức ‘GET’ và ‘POST’ sử dụng Flask:



* Bởi vì dữ liệu của ảnh trả về là dữ liệu dạng nhị phân nên chúng ta cần decode ảnh về dạng base64 để có thể đính hình ảnh vào HTML web



Ảnh được trả về dưới dạng base64

# CHƯƠNG IV: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## 4.1 Kết luận

Trong báo cáo này, em đã trình bày chi tiết quy trình xây dựng và huấn luyện mô hình YOLO (You Only Look Once) để nhận diện tình trạng hoa quả, cụ thể là phân loại hoa quả hư hại hay tươi tốt. Thông qua các bước chuẩn bị dữ liệu, tiền xử lý, thiết kế và tinh chỉnh mô hình, cũng như đánh giá hiệu suất, báo cáo đã chứng minh được hiệu quả của mô hình YOLO trong việc giải quyết bài toán nhận diện hình ảnh.

Kết quả thực nghiệm cho thấy mô hình YOLO có khả năng phát hiện và phân loại chính xác với độ chính xác cao. Các chỉ số đánh giá như mAP (mean Average Precision) và IoU (Intersection over Union) đều đạt được kết quả tốt, chứng tỏ mô hình có thể áp dụng vào thực tế với hiệu suất đáng tin cậy. Đặc biệt, thời gian suy luận nhanh của mô hình YOLO là một ưu điểm nổi bật, phù hợp cho các ứng dụng yêu cầu thời gian thực như trong dây chuyền kiểm tra chất lượng hoa quả tại các nhà máy chế biến.

Tuy nhiên, để mô hình hoạt động hiệu quả hơn trong môi trường thực tế, cần phải xem xét một số yếu tố như sự đa dạng của dữ liệu huấn luyện, khả năng xử lý các điều kiện ánh sáng và góc chụp khác nhau, và việc cập nhật mô hình theo thời gian với dữ liệu mới. Việc kết hợp mô hình YOLO với các kỹ thuật xử lý hình ảnh và học sâu khác cũng có thể là hướng nghiên cứu tiềm năng trong tương lai.

Tóm lại, việc áp dụng mô hình YOLO cho bài toán nhận diện hoa quả hư hại hay tươi tốt là một bước tiến quan trọng, mở ra nhiều cơ hội ứng dụng trong ngành nông nghiệp và công nghiệp thực phẩm, góp phần nâng cao chất lượng sản phẩm và hiệu quả sản xuất.

## 4.2 Hướng phát triển

Mặc dù mô hình YOLO đã cho thấy hiệu quả cao trong việc nhận diện hoa quả hư hại hay tươi tốt, vẫn còn nhiều hướng phát triển tiềm năng để nâng cao chất lượng và độ chính xác của hệ thống. Dưới đây là một số hướng phát triển chính:

1. **Tăng cường dữ liệu huấn luyện**: Mở rộng và đa dạng hóa tập dữ liệu huấn luyện là một bước quan trọng. Việc thu thập thêm dữ liệu từ nhiều nguồn khác nhau, bao gồm các loại hoa quả khác nhau, các điều kiện ánh sáng và môi trường khác nhau sẽ giúp mô hình trở nên linh hoạt và chính xác hơn.
2. **Tích hợp các phương pháp học sâu khác**: Kết hợp mô hình YOLO với các kỹ thuật học sâu tiên tiến như Transformer, GAN (Generative Adversarial Networks) hay các mô hình học sâu đa nhiệm để cải thiện khả năng nhận diện và phân loại.
3. **Tối ưu hóa mô hình**: Nghiên cứu và áp dụng các phương pháp tối ưu hóa mô hình như giảm kích thước mô hình, cải thiện tốc độ suy luận và giảm yêu cầu về tài nguyên tính toán sẽ giúp hệ thống trở nên nhẹ nhàng và hiệu quả hơn, đặc biệt trong các ứng dụng thời gian thực.
4. **Phát triển hệ thống giám sát liên tục**: Xây dựng hệ thống giám sát và cập nhật mô hình liên tục bằng cách thu thập và phân tích phản hồi từ người dùng, từ đó cập nhật mô hình với dữ liệu mới và cải thiện độ chính xác theo thời gian.
5. **Ứng dụng vào thực tiễn**: Mở rộng phạm vi ứng dụng của mô hình vào các lĩnh vực liên quan như giám sát chất lượng hoa quả trong dây chuyền sản xuất, kiểm tra hàng hóa tại các cửa hàng bán lẻ, hoặc ứng dụng trong nông nghiệp thông minh để theo dõi và quản lý cây trồng.
6. **Nghiên cứu tính khả thi và triển khai thực tế**: Tiến hành các nghiên cứu thử nghiệm và triển khai hệ thống trong môi trường thực tế để đánh giá tính khả thi và hiệu quả của mô hình. Qua đó, tìm ra các vấn đề tiềm ẩn và giải pháp khắc phục để hoàn thiện hệ thống.

Việc tiếp tục phát triển và cải tiến mô hình YOLO không chỉ giúp nâng cao độ chính xác và hiệu suất của hệ thống nhận diện hoa quả mà còn mở ra nhiều cơ hội ứng dụng mới, góp phần thúc đẩy sự phát triển của ngành công nghiệp thực phẩm và nông nghiệp.

# TÀI LIỆU THAM KHẢO

* + <https://docs.ultralytics.com/>
  + <https://yolov8.com/>
  + <https://www.youtube.com/watch?v=Mpsdby8zCTY>
  + <https://flask.palletsprojects.com/en/3.0.x/>
  + <https://www.youtube.com/watch?v=Z1RJmh_OqeA>