

**KHOA KỸ THUẬT VÀ CÔNG NGHỆ - ĐẠI HỌC HUẾ**

**BỘ MÔN THỰC TẬP DOANH NGHIỆP 2**

🙠🙟🕮🙝🙢



**BÁO CÁO**

**THỰC TẬP DOANH NGHIỆP**

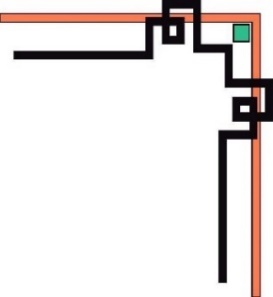
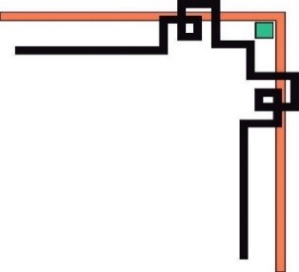
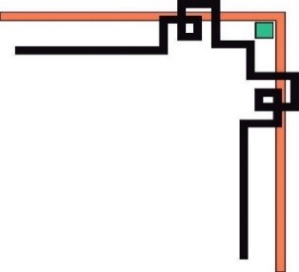
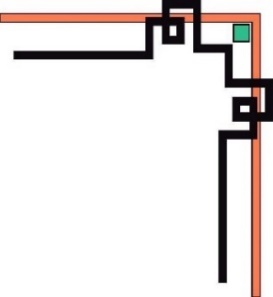
**NĂM HỌC 2023-2024**

**CHỦ ĐỀ THỰC TẬP:**

**DỰNG DEMO NHẬN DIỆN ĐÁM CHÁY BẲNG YOLO**

|  |
| --- |
| **Số phách**  *(Do hội đồng chấm thi ghi)* |

**Thừa Thiên Huế, ngày …tháng…năm.....**



**KHOA KỸ THUẬT VÀ CÔNG NGHỆ - ĐẠI HỌC HUẾ**

**BỘ MÔN THỰC TẬP DOANH NGHIỆP 2**

🙠🙟🕮🙝🙢



**BÁO CÁO**

**THỰC TẬP DOANH NGHIỆP**

**NĂM HỌC 2023-2024**

**CHỦ ĐỀ THỰC TẬP:**

**DỰNG DEMO NHẬN DIỆN ĐÁM CHÁY BẲNG YOLO**

**Giảng viên hướng dẫn: Hồ Quốc Dũng**

**Sinh viên thực hiện: Phan Thế Minh Châu**

**MSSV:20E1020067**

**Lớp: KHDL & TTNT - K1**

**Thừa Thiên Huế, ngày 19 tháng 01 năm 2024**

**ĐẠI HỌC HUẾ**

**KHOA KỸ THUẬT VÀ CÔNG NGHỆ**



**PHIẾU ĐÁNH GIÁ ĐỒ ÁN/TIỂU LUẬN/BÀI TẬP LỚN Học kỳ I, năm học 2023 - 2024**

|  |  |
| --- | --- |
| **Cán bộ chấm thi 1** | **Cán bộ chấm thi 2** |
| **Nhận xét:**  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  **Điểm đánh giá của CBCT1:**  Bằng số: .........................................  Bằng chữ: ....................................... | **Nhận xét:**  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  ..............................................................  **Điểm đánh giá của CBCT2:**  Bằng số: .........................................  Bằng chữ: ....................................... |

Điểm kết luận:

Bằng số:

Bằng chữ:

*Thừa Thiên Huế, ngày…tháng…năm 2023*

**Cán bộ chấm thi 1 Cán bộ chấm thi 2**

*(Ký và ghi rõ họ và tên) (Ký và ghi rõ họ và tên)*

**LỜI CẢM ƠN**

Trong suốt thời gian tám tuần thực tập là cơ hội cho nhóm em tổng hợp và hệ thống lại những kiến thức đã học, đồng thời kết hợp với thực tế để nâng cao kiến thức chuyên môn. Tuy chỉ có tám tuần thực tập, nhóm đã tiếp thu rất nhiều kiến thức thực tế. Từ đó nhóm em nhận thấy, việc cọ sát thực tế là vô cùng quan trọng, giúp sinh viên xây dựng nền tảng lý thuyết được học ở trường vững chắc hơn. Trong quá trình thực tập, từ bỡ ngỡ đến thiếu kinh nghiệm, nhóm đã gặp phải rất nhiều khó khăn. Tuy nhiên, nhóm đã nhận được rất nhiều sự quan tâm, giúp đỡ của thầy cô và công ty TNHH MTV khai thác dữ liệu số bData.

Nhóm xin chân thành gửi lời cảm ơn đến Thầy Cô ở Khoa Kỹ Thuật và Công Nghệ, Đại học Huế. Nhóm xin cảm ơn đặc biệt đến thầy Ts. Hồ Quốc Dũng trong thời gian qua đã tạo cơ hội cho nhóm em được trải nghiệm và làm việc thực tế tại công ty bData. Cảm ơn công ty bData đã tạo điều kiện cho nhóm tham gia thực tập ở công ty, học hỏi và khám phá nhiều kiến thức mới trong quá trình thực tập và phát triển kỹ năng trải nghiệm làm việc.Trong quá trình thực tập, hoàn thiện báo cáo thực tập em không tránh khỏi những sai sót, kính mong nhận được sự góp ý quý giá từ quý thầy cô cũng như anh, chị trong công ty.

Em xin chân thành cảm ơn!

# MỤC LỤC

[MỤC LỤC 1](#_Toc156428085)

[DANH MỤC HÌNH ẢNH 2](#_Toc156428086)

[LỜI MỞ ĐẦU 3](#_Toc156428087)

[1. Lý do chọn đề tài thực tập 3](#_Toc156428088)

[2. Mục tiêu thực tập 3](#_Toc156428089)

[3. Nhiệm vụ thực tập 3](#_Toc156428090)

[CHƯƠNG 1: TỔNG QUAN VỀ CÔNG TY TNHH MTV KHAI THÁC DỮ LIỆU SỐ BDATA 4](#_Toc156428091)

[1.1. Thông tin công ty bData 4](#_Toc156428092)

[1.2. Ngành nghề kinh doanh: 4](#_Toc156428093)

[CHƯƠNG 2: TỔNG QUAN LÝ THUYẾT 5](#_Toc156428094)

[2.1. Khái niệm về YOLO 5](#_Toc156428095)

[2.2. Kiến trúc mạng YOLO 5](#_Toc156428096)

[2.3. Nguyên lý hoạt động của mạng YOLO 6](#_Toc156428097)

[2.4. Các phiên bản của YOLO 7](#_Toc156428098)

[2.5. Cấu trúc tập dữ liệu YOLOv5 8](#_Toc156428099)

[CHƯƠNG 3: CHỌN DATASET ĐỂ HUẤN LUYỆN VÀ KIỂM THỬ 9](#_Toc156428100)

[3.1. Mô hình hệ thống 9](#_Toc156428101)

[3.2. Chuẩn bị dữ liệu 10](#_Toc156428102)

[3.3. Huấn luyện mô hình 11](#_Toc156428103)

[CHƯƠNG 4: KẾT QUẢ THỰC NGHIỆM 13](#_Toc156428104)

[CHƯƠNG 5: KẾT LUẬN 15](#_Toc156428105)

[TÀI LIỆU THAM KHẢO 16](#_Toc156428106)

# DANH MỤC HÌNH ẢNH

[Hình 1.1. Logo công ty bData 4](#_Toc156428107)

[Hình 2.1. Kiến trúc mạng của YOLO 5](#_Toc156428108)

[Hình 2.2. Mô hình hoạt động của mạng YOLO 6](#_Toc156428109)

[Hình 2.3. Định dạng File.txt gán nhãn của YOLOv5 8](#_Toc156428110)

[Hình 3.1. Mô hình hệ thống 9](#_Toc156428111)

[Hình 3.2. Bộ dữ liệu đám cháy 10](#_Toc156428112)

[Hình 3.3. Ảnh dữ liệu được gán nhãn trên Roboflow 10](#_Toc156428113)

[Hình 4.1. Biểu đồ kết quả training của model 13](#_Toc156428114)

[Hình 4.2. Kết quả nhận diện của tập test 13](#_Toc156428115)

[Hình 4.3. Kết quả nhận diện hình ảnh từ URL 14](#_Toc156428116)

# LỜI MỞ ĐẦU

## 1. Lý do chọn đề tài thực tập

Trong thời đại ngày nay, vấn đề an ninh và an toàn ngày càng trở nên quan trọng, đặc biệt là trong việc phát hiện và xử lý sự cố đám cháy. Đám cháy không chỉ gây thiệt hại về mặt tài sản mà còn đe dọa tính mạng con người. Theo thống kê của Cục Cảnh sát Phòng cháy chữa cháy và Cứu nạn cứu hộ, trong năm 2022, cả nước đã xảy ra 10.258 vụ cháy, làm chết 451 người, bị thương 1.144 người, thiệt hại tài sản ước tính 3.960 tỷ đồng. Để giảm thiểu thiệt hại do cháy nổ gây ra, việc phát hiện sớm đám cháy là vô cùng quan trọng. Hiện nay, sự phát triển nhanh chóng của cảm biến và công nghệ cũng như thuật toán thị giác máy tính, các phương pháp mới để phát hiện cháy đã được đề xuất. Tuy nhiên, các phương pháp này gặp phải một số hạn chế cần được giải quyết, đó là sự hiện diện của các vật thể giống lửa, tỷ lệ cảnh báo sai cao,...

Vì vậy, đồ án "Nhận Diện Đám Cháy bằng YOLOv5" để giải quyết vấn đề nhận diện đám cháy. Mục tiêu của đề tài là xây dựng một hệ thống nhận diện đám cháy linh hoạt và chính xác từ hình ảnh và video, góp phần vào công cuộc bảo vệ an toàn cộng đồng và tài sản.

## 2. Mục tiêu thực tập

Tham gia thực là cơ hội để nhóm học hỏi cách thức thực hiện công việc trong môi trường làm việc thực tế. Học hỏi và thực hành những kiến thức và kỹ năng trong quá trình thực tập. Xây dựng kinh nghiệm làm việc và tăng cơ hội tìm kiếm việc làm sau khi tốt nghiệp.

## 3. Nhiệm vụ thực tập

Có thể khám phá một ngành nghề hoặc lĩnh vực mới, thu được kinh nghiệm làm việc và liên kết, và tăng cơ hội được tuyển dụng sau khi tốt nghiệp. Để hoàn thành một nhiệm vụ thực tập, nhóm có thể cần phải làm các việc sau:

* Tìm hiểu về công ty mà nhóm sẽ thực tập.
* Hợp tác người hướng dẫn của nhóm để xác định mục tiêu, các nhiệm vụ cụ thể của nhóm và học hỏi được nhiều kiến thức trong quá trình thực tập.
* Ghi nhận và báo cáo lại công việc và các hoạt động đã thực hiện cho người hướng dẫn của nhóm.

# TỔNG QUAN VỀ CÔNG TY TNHH MTV KHAI THÁC DỮ LIỆU SỐ BDATA

## Thông tin công ty bData



Hình .. Logo công ty bData

Tên quốc tế: CÔNG TY TNHH MTV KHAI THÁC DỮ LIỆU SỐ

Tên viết tắt: CÔNG TY TNHH MTV KHAI THÁC DỮ LIỆU SỐ

Mã số thuế: 3301550690

Địa chỉ thuế: 29/19 Hoàng Diệu, Phường Tây Lộc, Thành phố Huế, Thừa Thiên Huế

Người đại diện: Nguyễn Hoài Nam

Điện thoại: 0988512889

Quản lý bởi: Cục thuế Tỉnh TT- Huế

Loại hình DN: Công ty trách nhiệm hữu hạn một thành viên

Tình trạng: Đang hoạt động (đã được cấp GCN ĐKT) [1]

## Ngành nghề kinh doanh:

* Xuất bản phần mềm.
* Lập trình máy vi tính.
* Tư vấn máy tính và quản trị hệ thống máy vi tính.
* Hoạt động dịch vụ công nghệ thông tin và dịch vụ khác liên quan đến máy tính.
* Xử lý dữ liệu, cho thuê và các hoạt động liên quan. [1]

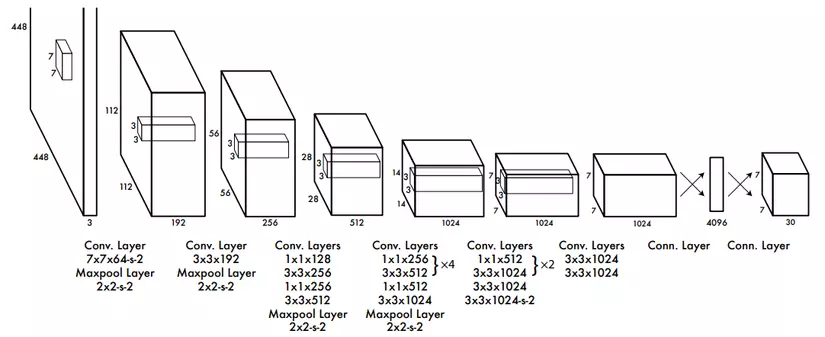
# TỔNG QUAN LÝ THUYẾT

## Khái niệm về YOLO

YOLO là một kiến trúc mạng CNN được sử dụng trong phát hiện, nhận dạng và phân loại đối tượng. Đối bài toán phân loại (Classification) chỉ có khả năng phân loại đối tượng bằng các dự đoán nhãn thì YOLO giải quyết bài toán phát hiện đối tượng (Object Detection), không chỉ có thể phát hiện nhiều đối tượng với nhiều nhãn khác nhau mà còn có thể xác định vị trí cụ thể của các đối tượng trong cùng một hình ảnh bằng các khung bao quanh đối tượng hình chữ nhật (Bounding Box). YOLO là viết tắt của cụm từ “You only look once” nói nên khả năng về tốc độ nhận dạng của mô hình này, YOLO được đánh giá là mô hình cho tốc độ nhận dạng nhanh nhất có khả năng nhận dạng theo thời gian thực. Kiến trúc YOLO được xây dựng từ các lớp tích chập (Convolution layers) để trích xuất ra các đặc trưng của đối tượng và các lớp kết nối đầy đủ (full connected layer) để dự đoán nhãn và vị trí của đối tượng . Dữ liệu đầu vào là các hình ảnh, mô hình sẽ dự đoán vị trí, kích thước và nhãn của các Bounding Box[2].

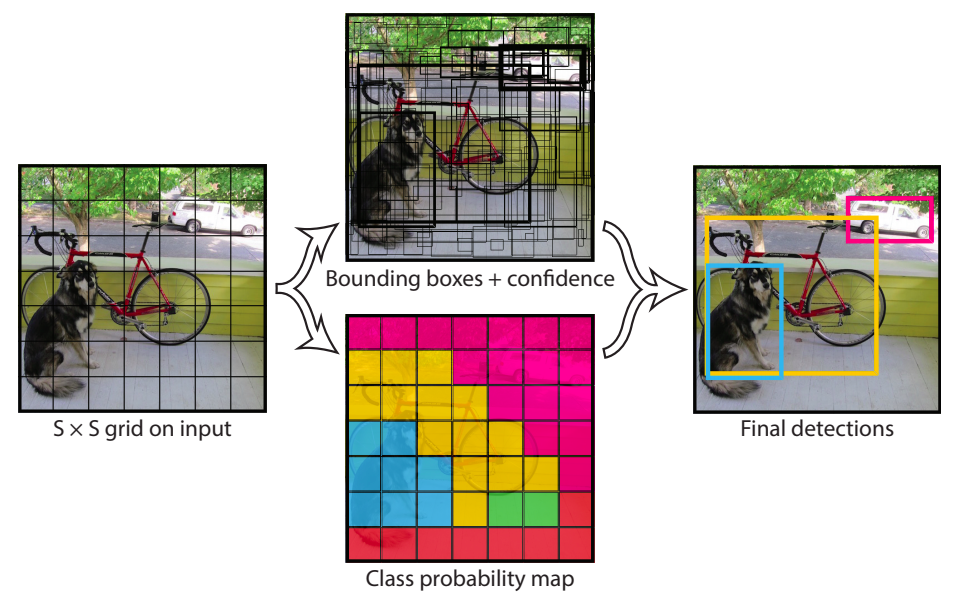
## 2.2. Kiến trúc mạng YOLO

YOLO có kiến trúc bao gồm phần trích xuất đặc trưng (Feature Extractor) và phát hiện đối tượng (Extra Layers). Phần trích xuất đặc trưng là các lớp convolution sẽ cho đầu ra là các bản đồ đặc trưng (Feature map). Phần phát hiện vật thể bao gồm các lớp kết nối đầy đủ (Fully connected layers) dùng để phát hiện, dự đoán nhãn và tọa độ Bounding Box của đối tượng trên các bản đồ đặc trưng (Feature map) của phần Feature Extractor. Hình ảnh cần nhận dạng sẽ được đi qua khâu trích xuất đặc trưng để có đầu ra ra là bản đồ đặc trưng (Feature map) với các kích thước khác nhau. Các bản đồ đặc trưng sau đó được đưa qua khâu phát hiện đối tượng (Extra Layers) để dự đoán ra các thông tinvề tên đối tượng, vị trí và kích thước của Bounding Box bao quanh đối tượng.



Hình .. Kiến trúc mạng của YOLO

Dữ liệu đầu vào là hình ảnh hoặc một frame được trích xuất từ Video. Ảnh đầu vào sau đó được chia thành mạng lưới có S x S ô (hay còn gọi là grid size). Thông thường grid size sẽ là 3 x 3, 7 x 7, 9 x 9, … [2]



Hình .. Mô hình hoạt động của mạng YOLO

## 2.3. Nguyên lý hoạt động của mạng YOLO

YOLO sẽ phân chia hình ảnh thành các ô vuông dạng lưới (grid) và xác định xem trong mỗi ô vuông liệu có tâm của vật thể cần xác định hay không. Nếu có, mô hình YOLO sẽ khoanh vùng đối tượng bằng hộp mốc (Anchor Box), sau khi có kết quả lọc chính xác thì sẽ xuất ra bounding box. Thuật toán sử dụng một mạng nơ ron đơn với đặc trưng có được từ các feature map của các lớp tích chập để dự đoán bounding box ở mỗi ô và xác suất loại đối tượng nào đang được chứa bên trong. Sau cùng, ta sẽ có rất nhiều bounding box được thuật toán đưa ra với kích thước khác nhau. Sử dụng thuật toán Non-Maxima Suppresstion (NMS) ta có thể loại được các hầu hết các bounding box là miền bao của cùng một đối tượng, có tỉ lệ thấp và giữ lại các bounding box có tỉ lệ khớp cao.

Đầu ra của YOLO khi dự đoán 1 Bounding Box là một vector gồm các thành phần gồm:

* Xác xuất dự báo có vật thể (P0) được định nghĩa Pr(Object)∗IOU(pred,truth)
* Tọa độ tâm kích thước chiều rộng, cao của Bouding Box (x, y, w, h),
* Các vector phân phối xác xuất dự báo của các classes.

Kích thước vector đầu ra được tính bằng (5 + số classes) ví dụ khi ta huấn luyện nhận dạng 50 đối tượng(classes) thì kích thước vector đầu ra dự đoán cho mỗ Bounding Box sẽ có kích thước là 55. Như vậy đầu ra của cả mô hình là một ma trận 3 chiều có kích thước S×S×B×(5+C) trong đó SxS là kích thước của các Feature map ở phần trích xuất đặc trưng, B là số lượng Bounding Box tương ứng là số Anchor Box và tại mỗi ô(cell) áp dụng 3 Anchor Box như vậy B = 3, C là số classes. [2]

## 2.4. Các phiên bản của YOLO

Mô hình YOLO hiện tại có 5 phiên bản:

* Phiên bản YOLOv1.
* Phiên bản YOLOv2.
* Phiên bản YOLOv3 do Joseph Redmon phát triển từ 2016.
* Phiên bản YOLOv4 do Alexey Bochkovskiy phát triển năm 2020.
* Phiên bản YOLOv5 do Glenn Jocher phát hành trong cùng năm 2020.

Khác với những phiên bản trước, YOLOv5 được phát triển dựa trên PyTorch thay vì DarkNet [5]. Đây là một ưu điểm lớn YOLOv5 vì PyTorch phổ biến hơn do đó sẽ có nhiều tài liệu và hướng dẫn để tham khảo về mô hình này. YOLOv5 đề xuất 5 phiên bản như sau:

* YOLOv5n - Nano version.
* YOLOv5s - Small version.
* YOLOv5m - Medium version.
* YOLOv5l - Large version.
* YOLOv5x - Extra-large version.

Trong bài báo cáo này, nhóm tôi sử dụng mô hình YOLOv5s để nhận diện đám cháy. [2]

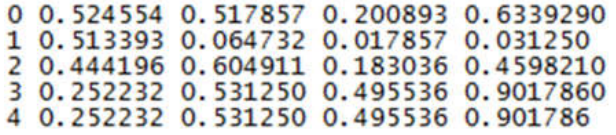
## 2.5. Cấu trúc tập dữ liệu YOLOv5

Một thư mục images chứa ảnh: trong đó có 2 thư mục train và val để chứa ảnh train và ảnh validate.

Một thư mục labels chứa nhãn (các file txt đó) và cũng có 2 thư mục tương tự như images.

Dữ liệu đầu vào ảnh của YOLOv5 theo format darknet với mỗi 1 file .txt sẽ cho 1 ảnh có chứa đối tượng là label, còn với những ảnh không có đối tượng thì bỏ qua. File .txt có định dạng như sau:

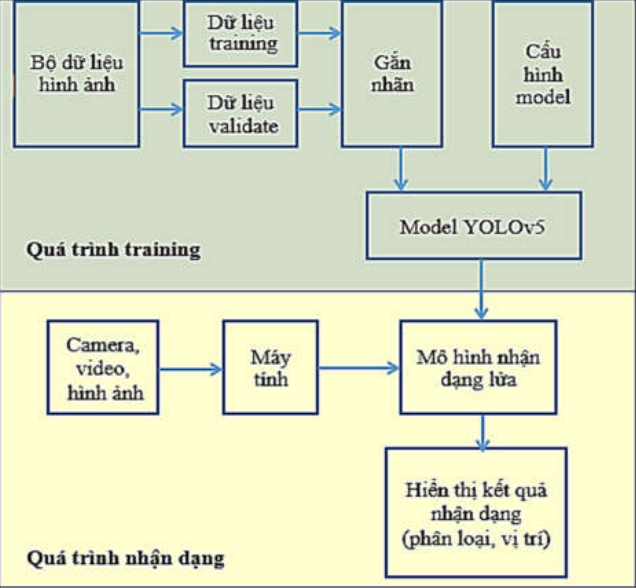
* Mỗi hàng sẽ là một đối tượng.
* Mỗi hàng sẽ có format như sau: class x\_center y\_center width height trong đó class x\_center y\_center là tọa độ x, y của tâm khung bao vật thể, width, height là chiều rộng và cao của khung bao vật thể.
* Toạ độ của các box sẽ được chuẩn hóa (từ 0-1) theo format xywh.
* Class sẽ bắt đầu từ 0 [2]



Hình .. Định dạng File.txt gán nhãn của YOLOv5

# CHƯƠNG 3: CHỌN DATASET ĐỂ HUẤN LUYỆN VÀ KIỂM THỬ

## 3.1. Mô hình hệ thống

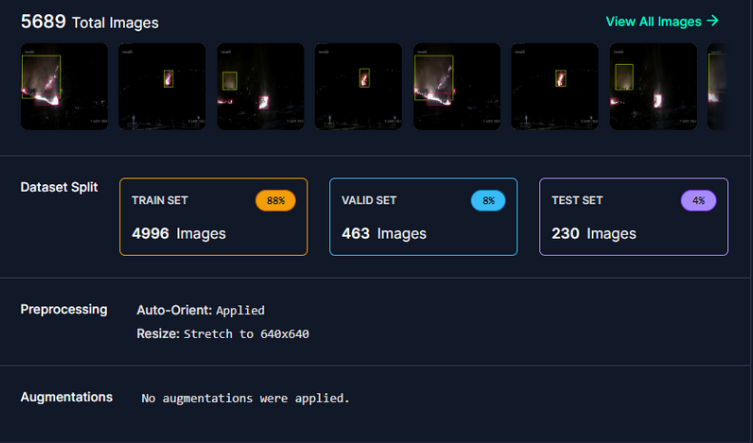


Hình 3.. Mô hình hệ thống

Bộ dữ liệu dataset được chia thành 3 tập dữ liệu bao gồm tập huấn luyện (training), tập xác minh (validate) và tập kiểm thử (test), các hình ảnh sau đó được tiến hành gán nhãn xác định vị trí các ngọn lửa.

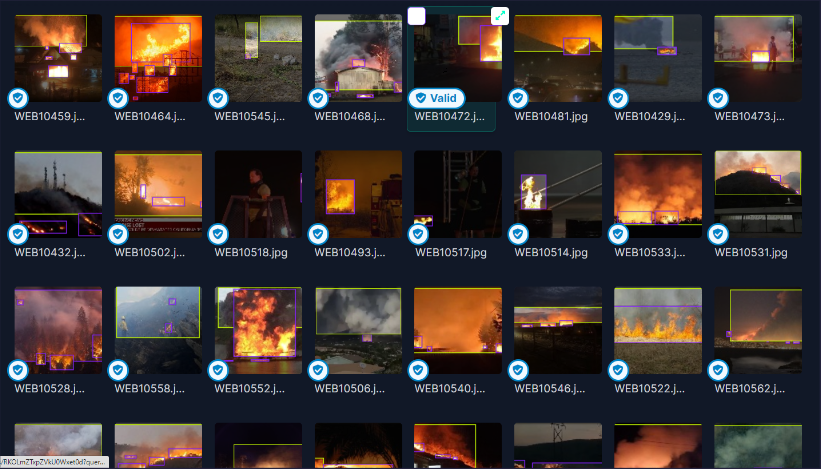
Tiến hành cấu hình các thông số cho Model, sử dụng tập dữ liệu huấn luyện, tập validate và tập test để huấn luyện cho mô hình. Mô hình sau khi được huấn luyện thành công bộ và được lưu dưới định dạng “best.pt”. Mô hình huấn luyện được sử dụng cho chương trình thực thi nhận dạng ngọn lửa với dữ liệu đầu vào lấy từ các hình ảnh, videos, camera. Kết quả quá trình nhận dạng sẽ hiển thị trực tiếp kết quả phân loại “fire” và vị trí của ngọn lửa trên màn hình.

## 3.2. Chuẩn bị dữ liệu



Hình 3.. Bộ dữ liệu đám cháy

Bộ dữ liệu được lấy từ link: <https://github.com/gaiasd/DFireDataset> gồm có 5689 ảnh và được gán nhãn dữ liệu trên Roboflow. Dữ liệu gán nhãn theo 2 đối tượng là lửa, khói và tổng số 5869 ảnh được chia 88% cho tập huấn luyện, 8% cho tập validate và 4% cho tập test.



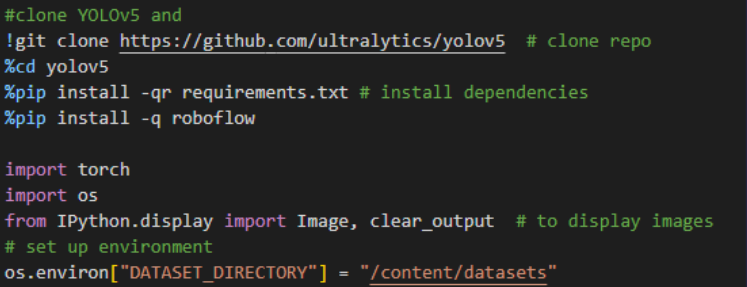
Hình 3.. Ảnh dữ liệu được gán nhãn trên Roboflow

## 3.3. Huấn luyện mô hình

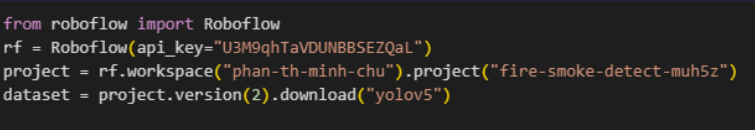
Trong quá trình huấn luyện mạng nơ ron sẽ tính toán lần lượt tất cả các ảnh của tệp dữ liệu đầu vào và sử dụng lặp lại các ảnh này nhiều lần, mục đích để tối ưu hàm mất mát. Quá trình tối ưu này sẽ giúp cho mạng nơ-ron tìm được bộ trọng số tốt nhất, giúp cho quá trình nhận dạng được chính xác nhất. Mô hình sau khi huấn luyện thành công bộ trọng số sẽ được lưu trữ trong file kết quả có định dạng “.pt”. Trong YOLOv5 ngoài việc bắt đầu quá trình huấn luyện mới từ đầu YOLOv5 còn cho phép người dùng sử dụng bộ trọng số có sẵn để tiếp tục quá trình huấn luyện để cho quá trình huấn luyện nhanh hơn thay vì phải huấn luyện từ đầu.

Quá trình huấn luyện mô hình sử dụng một lượng lớn tài nguyên của máy tính, nếu máy tính có cấu hình không đủ mạnh thời gian huấn luyện diễn ra rất dài. Ta có thể huấn luyện trên Google Colab để giảm thời gian huấn luyện đồng thời đảm bảo hiệu quả và chất lượng của mô hình.

Bước 1: Clone model và cài đặt các requirements:



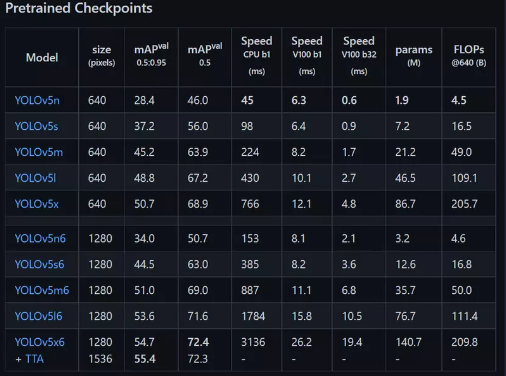
Bước 2: Clone dataset từ Roboflow:



Cấu trúc dataset được clone về:



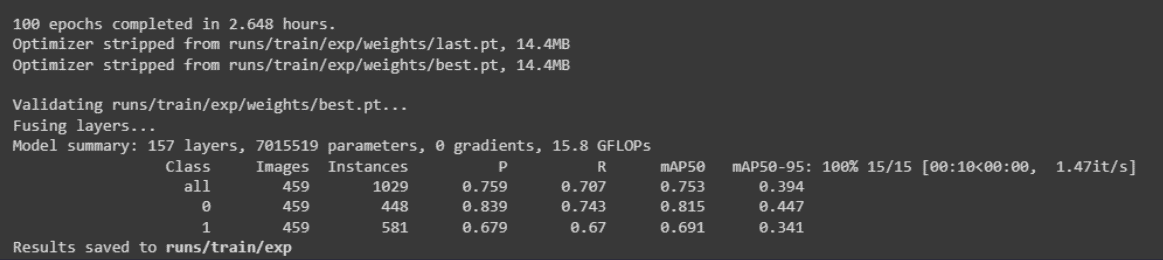
Bước 3: Chọn pretrained yolov5 với thông số phù hợp và tiến hành train bằng câu lệnh:



Chọn model yolov5s để tiến hành train.



Kết quả train:

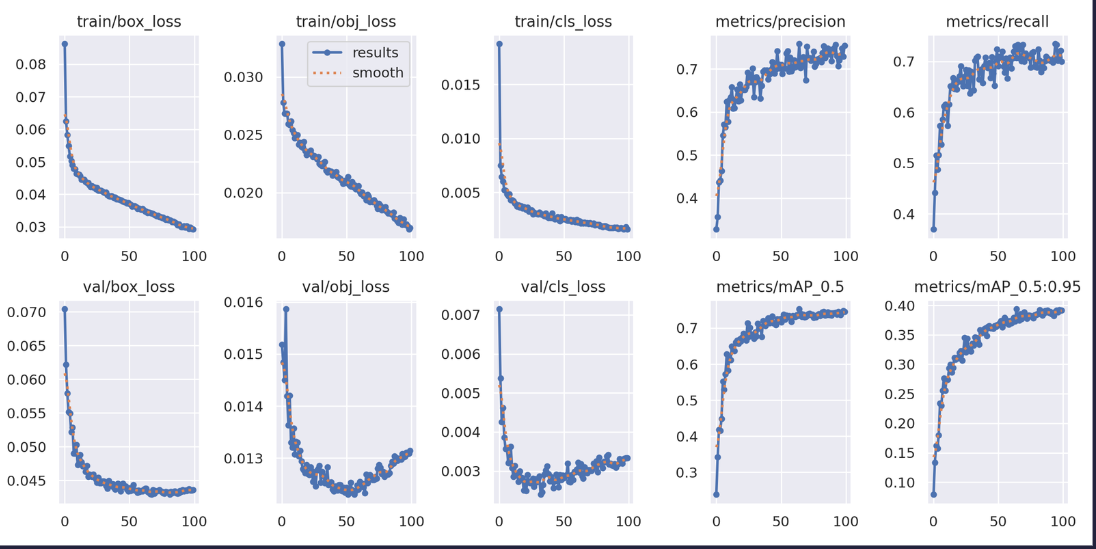


# CHƯƠNG 4: KẾT QUẢ THỰC NGHIỆM

Kết thúc quá trình huấn luyện có 2 bộ trọng số được mô hình lưu lại bao gồm:

* Bộ trọng số tốt nhất: best.pt
* Bộ trọng số cuối cùng của quá trình huấn luyện: last.pt

Trong đó bộ trọng số best.pt được sử dụng cho chương trình nhận dạng.



Hình 4.. Biểu đồ kết quả training của model

Biểu đồ hình 4.1 cho thấy, sau quá trình training 100 epoch độ chính xác trung bình dự đoán của training đạt 95%, độ mất mát (loss) 2.3%



Hình 4.. Kết quả nhận diện của tập test







Hình 4.. Kết quả nhận diện hình ảnh từ URL

Từ hình 4.2 và 4.3 cho thấy kết quả nhận diện đám cháy có độ chính xác cao và nhận diện rõ ràng lửa, khói một cách khác biệt không có sự trùng lẫn nhau.

# CHƯƠNG 5: KẾT LUẬN

Trong đề tài này, nhóm đã thực hiện nhận diện đám cháy trên bộ dữ liệu được lấy từ <https://github.com/gaiasd/DFireDataset> bằng việc sử dụng mô hình YOLOv5.

Kết quả cho thấy mô hình YOLOv5 rất phù hợp cho đề tài này vì cho ra hiệu năng tốt, độ chính xác cao và tốc độ nhận dang nhanh

Nhóm đề xuất mở rộng bộ dữ liệu và đồng thời nghiên cứu các mô hình nhận diện chi tiết hơn để lựa chọn mô hình phù hợp nhất. Ngoài ra, cần tập trung vào nghiên cứu kỹ thuật để giảm thời gian xử lý, nâng cao độ chính xác và nhận diện nhanh không bị trùng lặp, nhằm cải thiện hiệu suất hệ thống.

# TÀI LIỆU THAM KHẢO

[1] “3301550690 - CÔNG TY TNHH MỘT THÀNH VIÊN KHAI THÁC DỮ LIỆU SỐ &#128147; hosocongty.vn,” 3301550690 - CÔNG TY TNHH MỘT THÀNH VIÊN KHAI THÁC DỮ LIỆU SỐ &#128147; hosocongty.vn. Accessed: Jan. 17, 2024. [Online]. Available: https://hosocongty.vn/cong-ty-tnhh-mot-thanh-vien-khai-thac-du-lieu-so-com-505859.htm

[2] “Nâng cao hiệu năng phát hiện đám cháy sử dụng thị giác máy dựa trên mạng nơ-ron YOLOV5,” *jsthaui*, Oct. 2022, doi: 10.57001/huih5804.37.