

ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN

-----***-----



BÁO CÁO ĐỒ ÁN CUỐI KỲ

CS231.N21.KHCL

NHẬN DẠNG ĐỀN GIAO THÔNG

Sinh viên thực hiện:

20521938	Trần Phương Thảo
20521568	Nguyễn Hoàng Long

Hồ Chí Minh, 06 - 2023

[illegible]

MỤC LỤC

NHẬN XÉT CỦA GIẢNG VIÊN	2
-------------------------------	---

A. TỔNG QUAN BÀI TOÁN:	5
------------------------------	---

1. GIỚI THIỆU:	5
----------------------	---

2. MÔ TẢ:	5
-----------------	---

B. PHƯƠNG PHÁP:	6
-----------------------	---

1. TIÊU CHÍ CHỌN MÔ HÌNH:	6
---------------------------------	---

2. CÁCH THỰC HIỆN:	6
--------------------------	---

C. THỰC NGHIỆM:	8
-----------------------	---

1. DỮ LIỆU:	8
-------------------	---

1.1. MÔ TẢ:	8
-------------------	---

1.2. XỬ LÝ:	8
-------------------	---

1.3. THỐNG KÊ:	9
----------------------	---

2. ĐÁNH GIÁ MÔ HÌNH:	11
----------------------------	----

2.1. CẤU HÌNH THỰC HIỆN ĐÁNH GIÁ:	11
---	----

2.2. PRECISION:	11
-----------------------	----

2.3. RECALL:	11
--------------------	----

2.4. mAP:	12
-----------------	----

3. KẾT QUẢ THỰC NGHIỆM:	12
-------------------------------	----

D. CẢI THIỆN VÀ HƯỚNG PHÁT TRIỂN:	14
---	----

1. CẢI THIỆN:	14
2. HƯỚNG PHÁT TRIỂN:	14
TÀI LIỆU THAM KHẢO	15

A. TỔNG QUAN BÀI TOÁN:

1. GIỚI THIỆU:

Ngày nay, công nghệ phát hiện đèn giao thông có thể giúp người lái xe xác định trạng thái của đèn giao thông và đưa ra quyết định nhanh chóng theo trạng thái đã xác định của đèn giao thông. Điều này có thể làm giảm sự mất tập trung của người lái xe và ngăn chặn sự xuất hiện của các hành vi lái xe phi tiêu chuẩn và bất hợp pháp. Vì vậy, nghiên cứu về mô hình phát hiện và nhận dạng đèn giao thông với độ chính xác cao theo thời gian thực có ý nghĩa thực tiễn và triển vọng phát triển rộng lớn nhằm nâng cao an toàn giao thông đường bộ. Chính vì những lý do đó giúp mình có động lực để tìm hiểu về đề tài này.

Hệ thống phát hiện đèn giao thông thường sử dụng camera công nghiệp để thu thập thông tin về tình trạng đường xá. Do tính phức tạp và khả năng thay đổi của nền hình ảnh đèn giao thông trong các tình huống giao thông thực, đèn giao thông trong hình ảnh chiếm ít pixel hơn và cấu trúc đặc trưng của nó thưa thớt, điều này làm tăng độ khó nhận dạng của thuật toán. Do đó, việc nghiên cứu một thuật toán phát hiện mục tiêu nhỏ hiệu quả hơn để phát hiện đèn tín hiệu giao thông là rất quan trọng.

2. MÔ TẢ:

- Input:

Một video có độ dài bất kỳ với định dạng mp4 và độ phân giải tối thiểu 480p.

- Output:

Một video với thời lượng và nội dung giống như đầu vào với các bounding box chứa vùng ảnh có đèn giao thông ([link demo](#)).

B. PHƯƠNG PHÁP:

1. TIÊU CHÍ CHỌN MÔ HÌNH:

Sau khi nghiên cứu, tìm hiểu thì để phù hợp với môn học nhóm sử dụng những mô hình phổ biến, dễ tiếp cận như các mô hình YOLO hay RCNN. Nhưng sau khi tiếp cận thì nhóm chọn mô hình YOLO vì dựa trên:

- Thời gian thực thi (tốc độ).
- Độ chính xác.
- Hỗ trợ đầu vào là video.

2. CÁCH THỰC HIỆN:

Từ những tiêu chí trên, có thể thấy được mô hình YOLOv5 là 1 trong những state-of-the-art objects detector đáp ứng được. Và để so sánh và đánh giá về nhóm có sử dụng thêm mô hình YOLOv8.

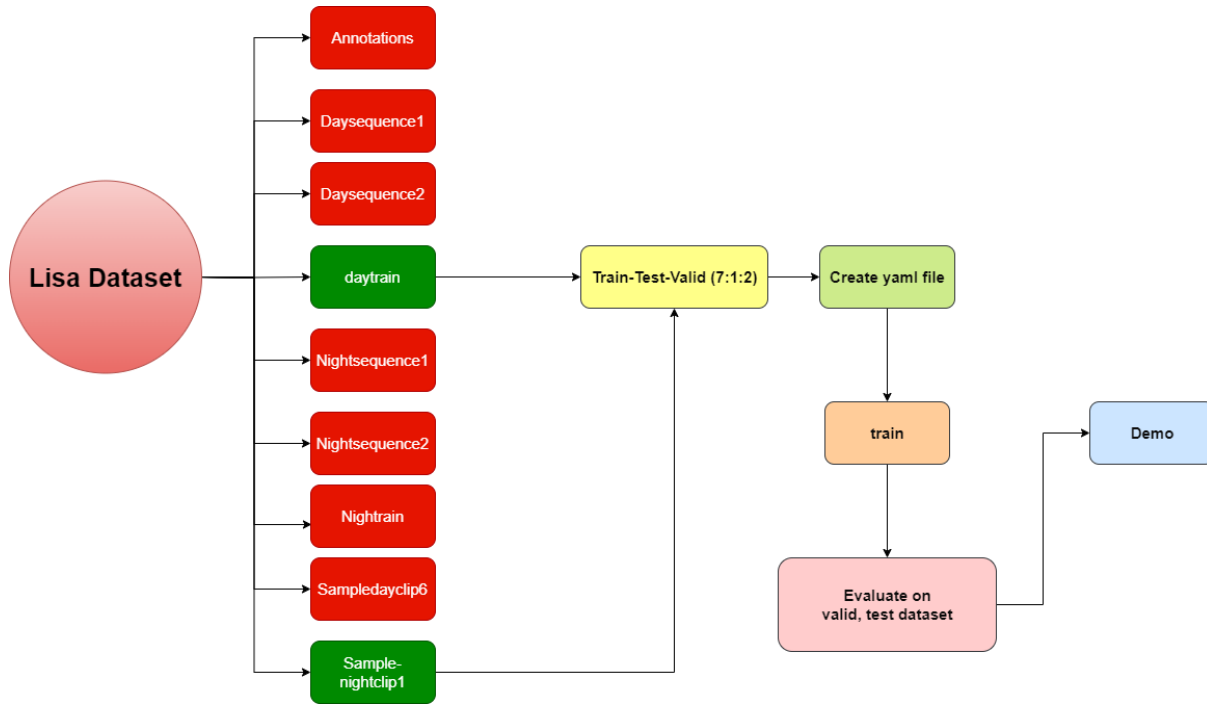
- YOLOv5

Mô hình YOLOv5 được ra đời vào năm 2020 được cải tiến từ mô hình YOLOv4 với nhiều cải tiến về kiến trúc mạng, quy trình huấn luyện, khả năng phát hiện và tốc độ xử lý, đồng thời hỗ trợ các phiên bản nhỏ hơn để sử dụng trên các thiết bị có tài nguyên hạn chế.

- YOLOv8

YOLOv8 được thiết kế để trở nên nhanh chóng, chính xác và dễ sử dụng, khiến nó trở thành lựa chọn tuyệt vời cho nhiều tác vụ phát hiện và theo dõi đối tượng, phân đoạn phiên bản, phân loại hình ảnh và ước tính tư thế.

Trong bài toán này ta tiến hành thực hiện các bước như sau:



Hình 1: Các bước thực hiện

Và trong quá trình training, ta thực hiện với các thông số sau:

	YOLOv5	YOLOv8
Image size	640	640
Epochs	20	20
Batch size	Auto	Auto
Weight	yolov5l	yolov8l

Bảng 1: Các parameters khi training.

Ở đây nhóm chọn YOLOv5l và YOLOv8l vì YOLOv5l có độ chính xác chỉ sau YOLOv5x mà số lượng params ít hơn và thời gian thực thi ít hơn YOLOv5x gần một nửa, YOLOv8l cũng tương tự.

C.THỰC NGHIỆM:

1. DỮ LIỆU:

1.1. MÔ TẢ:

- Dữ liệu thu thập từ Kaggle với bộ dữ liệu [LISA Traffic Light Dataset](#).
- Tập dữ liệu được thu thập bởi camera hành trình trong 4 ngày 2 đêm ở San Diego, California, USA.
- Gồm 9 folders Annotations, daySequence1, daySequence2, dayTrain, nightSequence1, nightSequence2, nightTrain, sample-dayClip6, sample-nightClip1.
- Kích thước ảnh trong tập dữ liệu này là 1280×960 . Và tập Annotations chứa các file txt chứa nội dung dưới dạng:

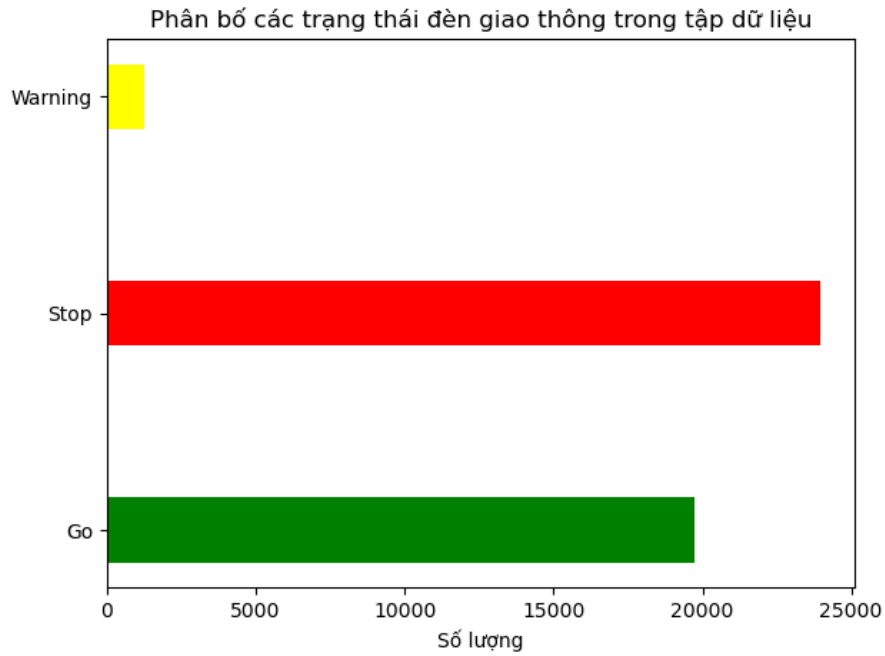
```
Filename;Annotation tag;Upper left corner X;Upper left corner  
Y;Lower right corner X;Lower right corner Y;Origin file;Origin  
frame number;Origin track;Origin track frame number
```

1.2. XỬ LÝ:

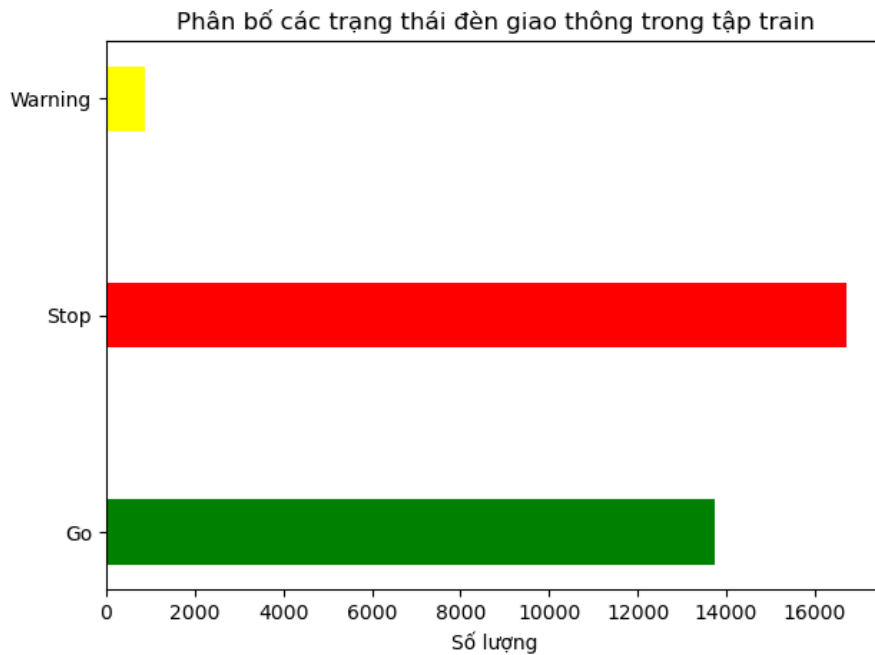
- Trong bài toán này, nhóm sử dụng 2 folder ảnh dayTrain và sample-nightClip1 và folder Annotations. Và chia theo tỷ lệ train-test-valid: 7-2-1.
- Do nhóm sử dụng 2 phương pháp YOLO nên phải tiến hành xử lý đưa file txt trong folder Annotations về dạng

```
Lable;X center;Y center;Width;Height
```

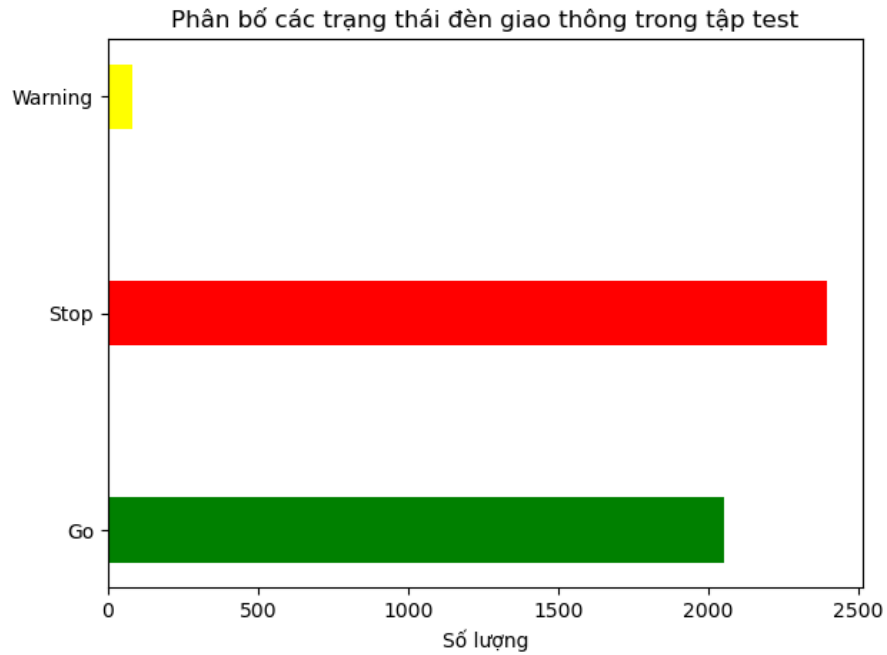

1.3. THỐNG KÊ:



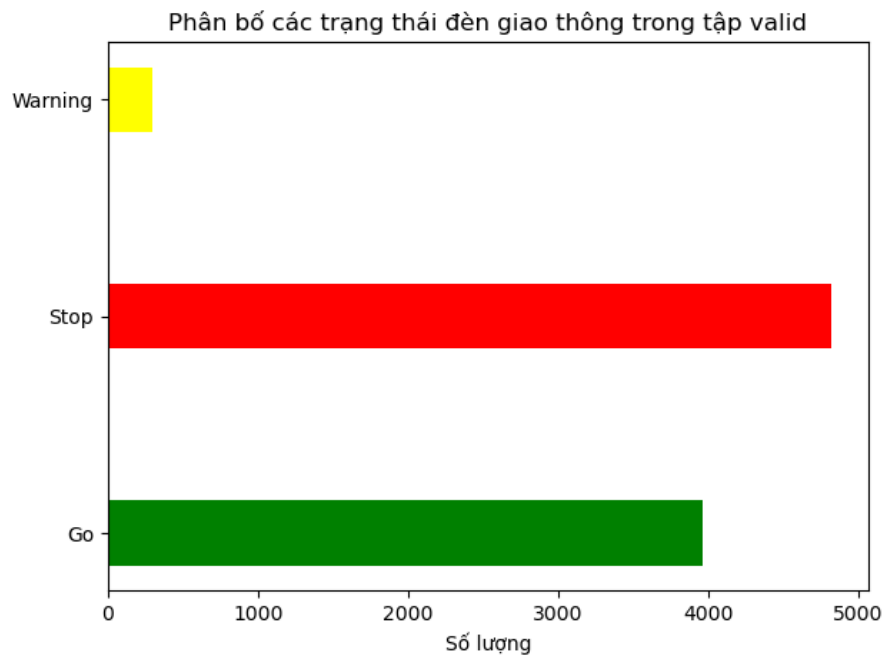
Hình 2: Phân bố các trạng thái đèn giao thông trong tập dữ liệu



Hình 3: Phân bố các trạng thái đèn giao thông trong tập train



Hình 4: Phân bố các trạng thái đèn giao thông trong tập test



Hình 5: Phân bố các trạng thái đèn giao thông trong tập valid

2. ĐÁNH GIÁ MÔ HÌNH:

2.1. CẤU HÌNH THỰC HIỆN ĐÁNH GIÁ:

Thực hiện đánh giá trên Google Colab với cấu hình GPU Tesla T4.

2.2. PRECISION:

Precision là tỷ lệ trường hợp bounding box có $\text{IoU} \geq \text{threshold}$ trong các bounding box được dự đoán.

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{TP}{\text{all detections}}$$

	YOLOv5	YOLOv8
all	0.929	0.946
go	0.963	0.951
warning	0.859	0.922
stop	0.966	0.965

Bảng 2: Precision của 2 mô hình YOLOv5 và YOLOv8

2.3. RECALL:

Recall đo lường tỷ lệ bounding box được dự đoán có $\text{IoU} \geq \text{threshold}$ trên tổng số ground truth box.

$$\text{Recall} = \frac{TP}{TP + FN} = \frac{TP}{\text{all ground truths}}$$

	YOLOv5	YOLOv8
all	0.947	0.904
go	0.961	0.907
warning	0.900	0.863
stop	0.981	0.943

Bảng 3: Recall của 2 mô hình YOLOv5 và YOLOv8

2.4. mAP:

mAP sẽ là trung bình cộng AP của tất cả các class.

	YOLOv5	YOLOv8
go	0.973	0.955
warning	0.910	0.909
stop	0.984	0.979

Bảng 4: mAP50 của 2 mô hình YOLOv5 và YOLOv8

	YOLOv5	YOLOv8
go	0.794	0.805
warning	0.52	0.534
stop	0.847	0.855

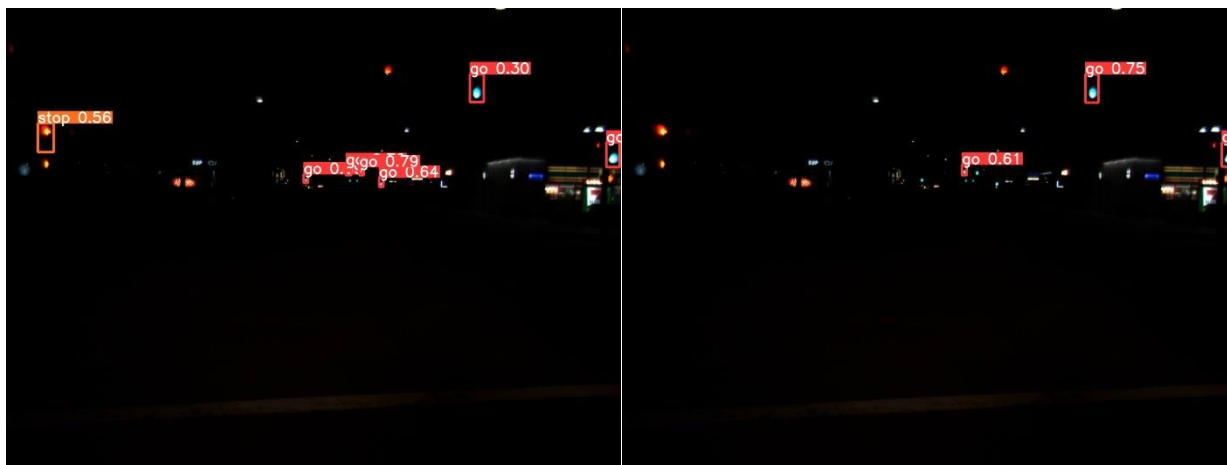
Bảng 5: mAP50-95 của 2 mô hình YOLOv5 và YOLOv8

3. KẾT QUẢ THỰC NGHIỆM:

Các kết quả mô hình dự đoán:



Hình 6: Kết quả mô hình dự đoán YOLOv5 (bên trái) và YOLOv8 (bên phải) với ảnh ban ngày.



Hình 7: Kết quả mô hình dự đoán YOLOv5 (bên trái) và YOLOv8 (bên phải) với ảnh ban đêm.



Hình 8: Kết quả mô hình dự đoán YOLOv5 (bên trái) và YOLOv8 (bên phải) với ảnh chụp cận

Nhận xét:

- Cả 2 mô hình đều có độ chính xác khá cao và các bounding box bao khá sát vào các đèn giao thông mô hình dự đoán. Tuy nhiên về độ phủ ta có thể thấy rằng YOLOv5 lại trội hơn so với YOLOv8.
- Do đây là bài toán dựa trên tập dữ liệu quay bằng camera hành trình nên dữ liệu về đèn giao thông ở trạng thái ‘warning’ khá ít dẫn đến việc cả 2 mô hình bỏ sót đi khá nhiều object này.
-

D.CẢI THIỆN VÀ HƯỚNG PHÁT TRIỂN:

1. CẢI THIỆN:

- Thu thập thêm dữ liệu, chủ yếu là đèn giao thông màu vàng và thu thập thêm ảnh nền cũng như những ảnh về đèn xe và đèn đường để tránh mô hình nhầm lẫn.
- Sử dụng thêm nhiều mô hình để chọn ra mô hình phù hợp nhất với bài toán.

2. HƯỚNG PHÁT TRIỂN:

- Áp dụng vào xử lý real-time.
- Ứng dụng trong phát triển xe tự hành.

TÀI LIỆU THAM KHẢO

- [1] Lisa Traffic Light Dataset - <https://www.kaggle.com/datasets/mbornoe/lisa-traffic-light-dataset>
- [2] YOLOv5 - <https://github.com/ultralytics/YOLOv5>
- [3] YOLOv8 - <https://github.com/ultralytics/ultralytics>
- [4] Evolution of YOLO object detection model from v5 to v8 –
<https://www.labellerr.com/blog/evolution-of-YOLO-object-detection-model-from-v5-to-v8/>