**HỘI NGHỊ SINH VIÊN NGHIÊN CỨU KHOA HỌC HUTECH NĂM 2024**

**Ứng dụng mô hình YOLOv8 trong nhận diện biển báo giao thông**

**Huỳnh Long1, Nguyễn Văn Đạt1\*, Nguyễn Chấn Huy1**

1  Viện Kỹ Thuật, Trường Đại học Công nghệ Thành phố Hồ Chí Minh.

\*Tác giả liên hệ: (Điện thoại: 0337141342; Email: datnvan.021504@gmail.com).

**Giảng viên hướng dẫn:** Thầy TS. Phạm Quốc Thiện

**TÓM TẮT**

Biển báo giao thông và các biển khác như biển cấm đỗ xe, biển cấm đi ngược chiều, v.v., là thành phần quan trọng khi lưu thông trên đường và đã trở nên rất quan trọng trong xe tự hành và xe tự lái vì nó giúp hệ thống nhận diện để đưa ra các lệnh điều khiển và tuân thủ các quy tắc giao thông cơ bản. Ngoài ra, nó còn giúp điều hướng các tuyến đường dựa trên các biển báo, từ đó tạo ra trải nghiệm lái xe an toàn hơn cho người lái. Đã có rất nhiều mô hình nhận diện mới xuất hiện trong những năm gần đây liên quan đến vấn đề này. Trong bài báo này, nghiên cứu đã sử dụng hệ thống phát hiện đối tượng YOLOv8 mới để giúp phát hiện biển báo giao thông vì nó nhanh hơn và chính xác hơn so với các phiên bản trước đó. Để cải thiện thuật toán, bài báo này đã sử dụng một bộ dữ liệu bao gồm ảnh của các biển báo giao thông được chụp ở các góc độ và cường độ ánh sáng khác nhau, để huấn luyện dữ liệu cho mô hình, từ đó cho thấy các chỉ số hiệu suất khi chỉ huấn luyện với 50 epochs với mAP5095, mAP50, precision, recall, F1 lần lượt là 68%, 82%, 75%, 82%, 78%.

Từ khóa : YOLOv8, biển báo giao thông, xe tự hành.

**I. Giới thiệu**

Hệ thống đường bộ là phần quan trọng trong cuộc sống và mỗi ngày có rất nhiều phương tiện lưu thông trên đường và để tránh ùn tắc giao thông, tránh các sự cố xảy ra khi lưu thông cũng như điều hướng phương tiện khi tham gia giao thông nên hệ thống biển báo giao thông, đèn giao thông,.. đã ra đời để thực hiện nhiệm vụ đó.

Biển báo giao thông rất quan trọng để duy trì an toàn và phối hợp trên đường khi lái xe. Có rất nhiều biển báo giao thông, và mỗi biển báo đại diện cho nhiệm vụ khác biệt của nó so với những biển khác, ví dụ như biển báo cấm đỗ xe, biển dừng và cấm ngược chiều đến biển báo cảnh báo rẽ trái hay rẽ phải, mỗi biển đều có một vai trò rất quan trọng trong an toàn giao thông. Trước đây, chỉ cần dạy và huấn luyện tài xế về các biển báo để họ có thể tuân theo những biển báo đó và duy trì các quy định để đảm bảo lái xe ổn định, an toàn.

Nhưng với sự tăng trưởng mạnh mẽ của công nghệ hiện đại thì hệ thống xe tự lái đang ngày càng phát triển. Và để đảm bảo an toàn cho người dùng khi sử dụng xe tự lái thì nhiệm vụ nhận diện phân loại biển báo một cách nhanh chóng, chính xác là một trong những nhiệm vụ được ưu tiên hàng đầu, điều này đòi hỏi việc nghiên cứu một cách liên tục.

Có rất nhiều loại biển báo khác nhau, mỗi loại đại diện cho một ý nghĩa nhất định nên nhiệm vụ phát hiện và phân loại biển báo giao thông là hai nhiệm vụ tối quan trọng để có thể nhận diện được biển báo. Trong những năm gần đây, việc sử dụng các mô hình học sâu và CNN đã tăng lên, nó giúp tăng cường và cải thiện quá trình của các mô hình phát hiện đối tượng để giải quyết vấn đề này trong thời gian thực vì CNN có khả năng trích xuất các đặc trưng hình ảnh và phân loại hình ảnh với độ chính xác. Và bài báo này sử dụng mô hình YOLOv8 là phiên bản mới của YOLO [1] nhằm thực hiện nhiệm vụ nhận diện biển báo theo thời gian thực.

**II. Tổng quan**

Hiện tại đã có rất nhiều mô hình học sâu được dùng để nhận diện cũng như phân loại vật thể như Faster RCNN, sử dụng một mạng đề xuất khu vực và loại bỏ việc sử dụng tìm kiếm lựa chọn. Điều này giúp đạt được phát hiện đối tượng gần như thời gian thực bằng cách cho phép mạng học các đề xuất khu vực. Tuy nhiên, sự dư thừa trong các giai đoạn phát hiện tiếp theo vẫn là một vấn đề đáng quan tâm.

Vào năm 2017, Feature Pyramid Networks được đề xuất. Nếu đi sâu vào Faster RCNN, thấy rằng nó chủ yếu không thể bắt được các đối tượng nhỏ trong hình ảnh. Để giải quyết điều này, có thể sử dụng một kim tự tháp hình ảnh đơn giản để chia tỷ lệ hình ảnh thành các kích cỡ khác nhau và gửi nó đến mạng.

Tất cả các thuật toán được đề cập ở trên chỉ phát hiện đối tượng cho một khu vực cục bộ trong hình ảnh và không hoạt động trên toàn bộ hình ảnh.

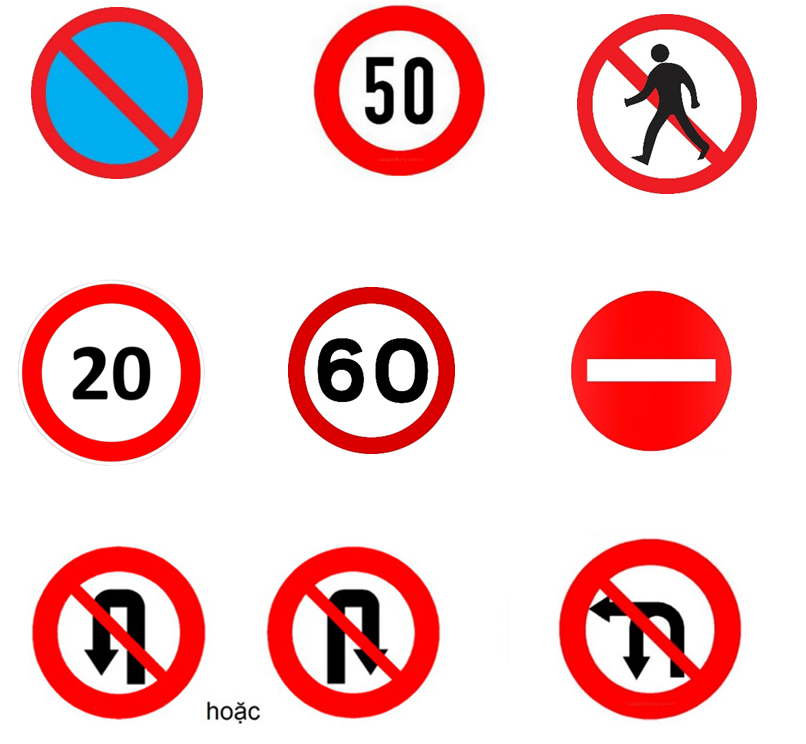
Các thuật toán YOLO [1] là thuật toán object detection nên mục tiêu của mô hình không chỉ là dự báo nhãn cho vật thể như các bài toán classification mà nó còn xác định location của vật thể. Do đó YOLO có thể phát hiện được nhiều vật thể có nhãn khác nhau trong một bức ảnh thay vì chỉ phân loại duy nhất một nhãn cho một bức ảnh.

Hình ảnh thu được được cắt thành các khu vực vuông với kích thước n\*n và trích xuất các đặc trưng từ nó. Các đặc trưng của tất cả các loại nhãn khác nhau được học từ các dữ liệu khác nhau. Khu vực mà dữ liệu được thu thập, khu vực mục tiêu ngang và dọc của nó được xác định và tất cả các thông số cần thiết được xác định để dự đoán đối tượng [2].

Giới thiệu về YOLO: YOLO [1] là một thuật toán phát hiện đối tượng hiện đại, được giới thiệu vào năm 2015 bởi Joseph Redmon, Santosh Divvala, Ross Girshick và Ali Farhadi trong bài báo nghiên cứu “You Only Look Once: Unified, Real-Time Object Detection”.

YOLO đưa ra vấn đề phát hiện đối tượng như một vấn đề hồi quy thay vì một nhiệm vụ phân loại. Thay vì tách không gian của khung viên bao xung quanh vật thể được nhận diện, YOLO liên kết xác suất với mỗi hình ảnh được phát hiện bằng cách sử dụng một mạng nơron tích chập (CNN) duy nhất.

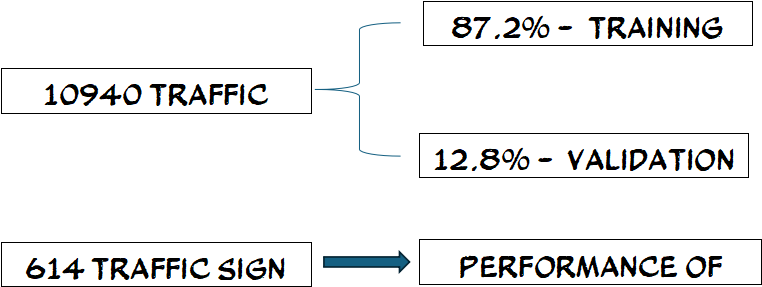
Mục tiêu sử dụng YOLO trong dự án: Mục tiêu sử dụng YOLO trong dự án là nhận diện biển báo giao thông để xe tự động đưa ra tín hiệu điều khiển khi chạy trên đường. Điều này giúp người dùng kiểm soát sự di chuyển của xe một cách an toàn.

Phạm vi dự án: Sử dụng mô hình YOLOv8 với model yolov8n để tiến hành nhận diện biển báo giao thông. Mô hình YOLOv8 trong dự án này có thể nhận diện được 58 biển báo giao thông. Dưới đây là một số trong 58 biển báo mà mô hình nhận diện được.

***Hình 1:*** *Một số biển báo mà mô hình có thể nhận diện được*

**III. Phương pháp**

Nguồn : hiện nay chủ đề nhận diện biển báo giao thông khá là phổ biến nên khá nhiều nguồn tham khảo để có thể tải dataset cho đề tài nhận diện biển báo giao thông. Và dữ liệu để thực hiện nhiệm vụ nhận diện biển báo giao thông cho dự án xe tự hành được tải từ trang Kaggle với định dạng YOLOv8 để tiến hành huấn luyện dữ liệu.



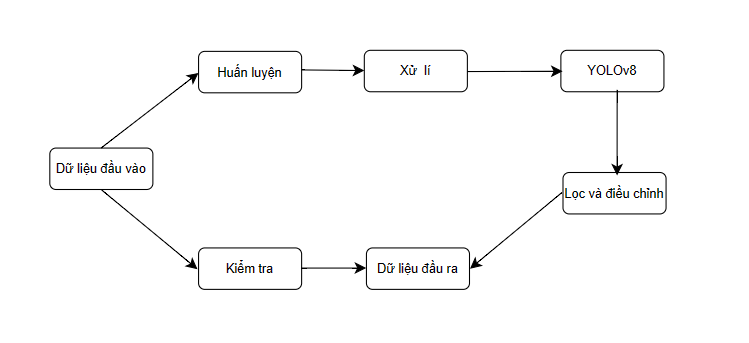
***Hình 2:*** *Sơ đồ diễn tả dữ liệu phân chia dataset*

Trong tệp dataset trên có tổng cộng 10940 ảnh (ảnh của 58 biển báo giao thông) dùng để huấn luyện và đánh giá hiệu suất huấn luyện của model, 9540 ảnh tương đương với khoảng 87,2% số ảnh trong số 10940 dùng để huấn luyện model và 786 ảnh tương đương với khoảng 12,8% dùng để đánh giá, và dùng 614 ảnh khác không có trong 10940 ảnh dùng để huấn luyện và đánh giá để kiểm thử mô hình YOLOv8.

Tệp dataset của được xây dựng trên nền tảng Roboflow Universe [4], trên nền tảng này khi xây dựng một tập dataset thì khi tải ảnh lên thì Roboflow sẽ hỗ trợ làm tất cả các ảnh sẽ có chung một định dạng và có thể gán nhãn trong nền tảng của Roboflow Anotate [5], và kích thước của hình ảnh dành cho huấn luyện là 640x640. Và cuối là xuất tệp dataset sang định dạng YOLOv8 để có thể tiến hành huấn luyện dữ liệu.

Trong bài báo này sử dụng chủ yếu ngôn ngữ Python [6] để tiến hành huấn luyện mô hình. Thông số cho quá trình huấn luyện để thực hiện nhiệm vụ nhận diện biển báo giao thông của dự án xe tự hành : epochs = 50, device = 0, batch = 2, workers = 2, model = yolov8n. Epochs là thuật ngữ đề cập đến số lần mà toàn bộ tập dữ liệu huấn luyện được sử dụng để đưa vào mạng nơron để cải thiện các trọng số của mô hình, device là thiết bị mà mô hình dùng để tiến hành huấn luyện ở đây là GPU của máy tính người dùng, batch là thông số quan trọng để cải thiện hiệu suất và ổn định mô hình, nó là tập hợp các mẫu dữ liệu được sử dụng để cập nhật các trọng số của mạng nơron trong mỗi lần lặp lại của quá trình huấn luyện, worker là thuật ngữ để ám chỉ các tiến trình hoặc luồng độc lập được sử dụng để xử lý các tác vụ nhất định. Cụ thể, khi sử dụng frame work Pytorch, việc sử dụng nhiều worker có thể làm tăng tốc độ huấn luyện bằng cách chia nhỏ các tác vụ và phân phối chúng đến các worker khác nhau để xử lý song song.

Từ tập dữ liệu đào tạo, tất cả các đặc điểm cần thiết được trích xuất sau đó được huấn luyện bằng thuật toán YOLOv8. Sau khi được huấn luyện, dữ liệu sau đó được điều chỉnh và lọc để loại bỏ một số điểm không chính xác và sau đó kiểm tra nó với bộ dữ liệu thử nghiệm để có được đầu ra. Hình 3 [7] giải thích toàn bộ kiến trúc của kiến trúc được đề xuất.

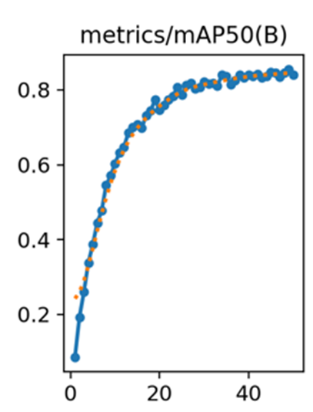
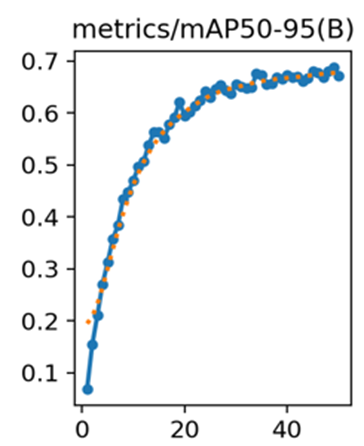
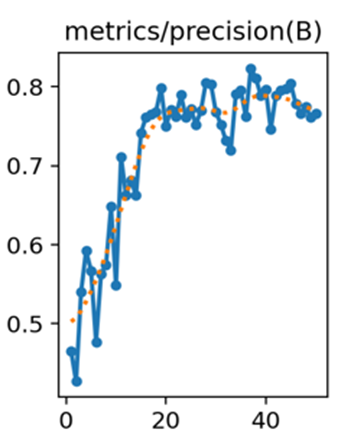
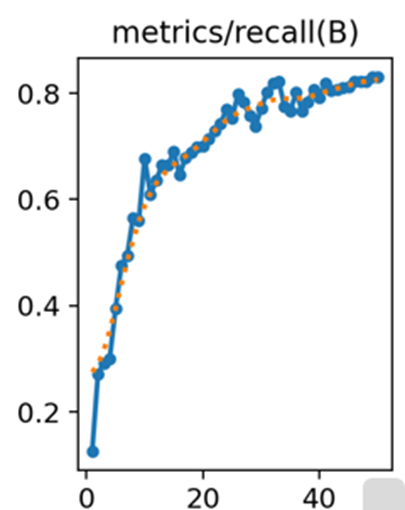
**

***Hình 3:*** *Sơ đồ diễn tả cấu trúc được đề xuất*

Yolov8 bao gồm một mạng nơ-ron tích chập (CNN) chia hình ảnh thành lưới và xác định các hộp giới hạn và lớp cho các đối tượng trong tất cả các ô lưới tại một thời điểm. Nó sử dụng một hộp neo để cải thiện độ chính xác cho hình ảnh có kích thước và hình dạng khác nhau và kết hợp các phương pháp khác nhau như mạng kim tự tháp để nâng cao hiệu suất.

**IV. Kết quả và thảo luận**

Dữ liệu có tổng cộng 58 biển báo với 10940 ảnh trong đó sẽ chia ra 87,2% để huấn luyện và 12,8% để đánh giá, 614 ảnh ngoài để kiểm thử. Mô hình YOLOv8 trong bài báo này được huấn luyện trên thiết bị laptop với 11th Gen Intel(R) Core(TM) i5-11400H @ 2.70GHz 2.69 GHz CPU, 16GB ram, NVIDIA GeForce RTX 3050 GPU. Dưới đây là các thông số kết quả của quá trình huấn luyện sau 50 epochs.



(b)

)

(a)

(c)

(d)

***Hình 4:*** *Biểu đồ biểu thị các chỉ số kết quả huấn luyện:*

*(a) mAP50, (b) mAP50-95, (c) precision và (d) recall*

Kết quả Huấn Luyện: Các biểu đồ cho thấy kết quả sau khi huấn luyện mô hình YOLOv8. Mô hình này không sử dụng tham số “accuracy” truyền thống mà thay vào đó dùng mAP50 và mAP5095 để đánh giá độ tin cậy của kết quả huấn luyện.

IoU (Intersection over Union): Đo lường sự chồng chéo giữa hộp giới hạn vật thể dự đoán và hộp giới hạn thực tế, là chỉ số cơ bản đánh giá độ chính xác vị trí đối tượng.

AP (Average Precision): Tính toán diện tích dưới đường cong của Precision và Recall, cung cấp thông tin về cách mô hình cân bằng giữa độ chính xác và độ phủ.

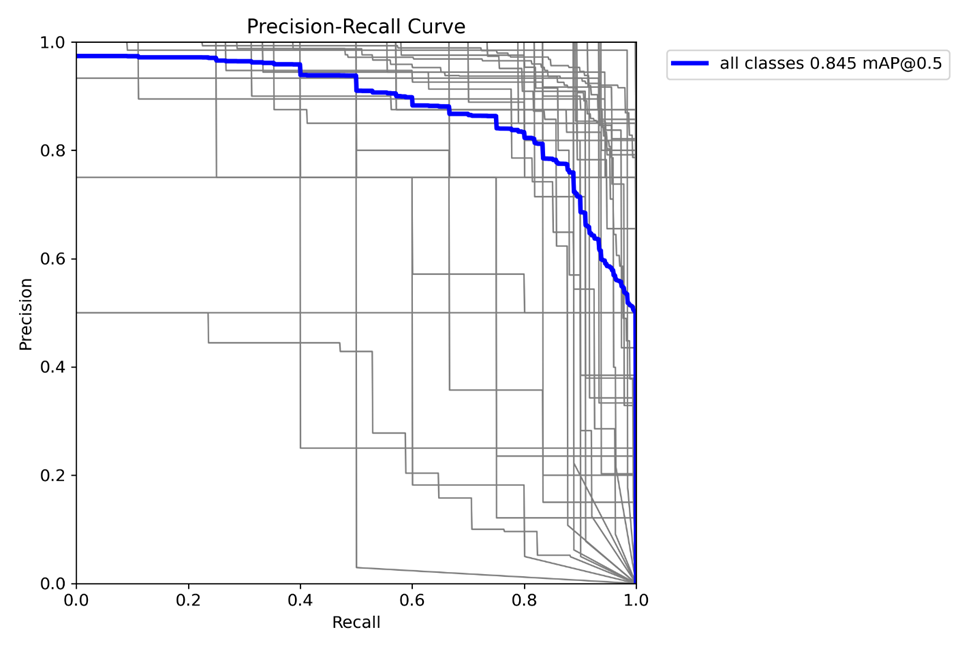
mAP (Mean Average Precision): Tính trung bình các giá trị AP qua nhiều lớp đối tượng, đánh giá hiệu suất tổng thể của mô hình trong phát hiện đa lớp đối tượng.

mAP50-95: Trung bình của mAP tại các ngưỡng IoU từ 50% đến 95%, cung cấp cái nhìn tổng quan về hiệu suất mô hình qua các cấp độ khác nhau của độ khó.

Precision và Recall: Hai chỉ số này đánh giá độ chính xác của dự đoán và khả năng phát hiện vật thể, giúp xác định độ tin cậy của dự đoán và hướng phát triển để cải thiện mô hình.

Chỉ số F1 được tính bằng công thức:

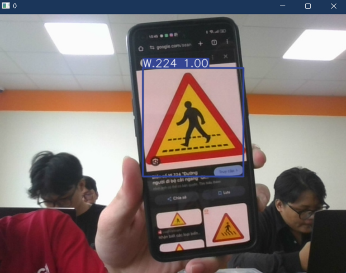
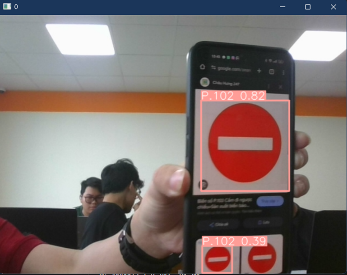
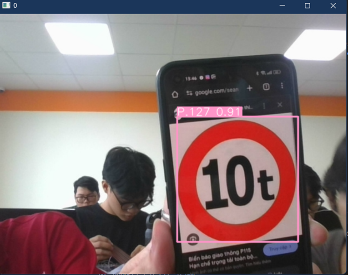


Nó là một phép đo tổng quát về khả năng của mô hình trong việc phát hiện và định vị vật thể trong hình ảnh. Mục tiêu khi huấn luyện mô hình YOLOv8 là tối ưu hóa chỉ số F1 để đảm bảo mô hình đạt được sự cân bằng giữa precision và recall. Và từ những biểu đồ trên có thể thấy kết quả của mô hình YOLOv8 khá ổn khi chỉ huấn luyện với 50 epochs với mAP5095, mAP50, precision, recall, F1 lần lượt là 68%, 82%, 75%, 82%, 78% . Và dưới đây là hình ảnh mô tả mối quan hệ giữa -precision và recall:

***Hình 5:*** *Sơ đồ mối tương quan giữa hai thông số precision và recall*

Từ Hình 8 có thể thấy rằng chỉ số recall càng thấp thì chỉ số precision càng cao và chỉ số recall càng cao thì chỉ số precision càng giảm điều đồng nghĩa với việc khi nhận diện càng nhiều vật thể thì độ chính xác sẽ giảm xuống nhưng vẫn đủ để có thể tin tưởng vào kết quả nhận diện và khi nhận diện càng ít vật thể thì độ chính xác của việc nhận diện càng cao.

Để thấy được độ chính xác và độ tin cậy của mô hình trong thực tế thì dưới đây là một số hình ảnh nhận diện từ thực tế để có thể đánh giá một cách chính xác hơn.



(b)

)

(a)

)

)

(c)

)

(d)

)

***Hình 6:*** *Kết quả nhận diện hình ảnh các biển báo:*

*(a) cảnh báo có người đi bộ, (b) cấm đi ngược chiều,*

*(c) giới hạn 10 tấn và (d) giới hạn tốc độ 50*

Có thể thấy rằng khả năng nhận diện của mô hình YOLOv8 vẫn ổn khi chỉ với 50 epochs nhưng tất nhiên là không thể chính xác 100% đúng tất cả các biển báo như hai hình ảnh Hình 10 và Hình 10 khi chỉ huấn luyện với khoảng hơn 10000 ảnh và 50 epochs. Tuy nhiên, vẫn không thể phủ nhận khả năng phát hiện và phân loại vật thể của YOLOv8 cụ thể hơn là nhận diện biển báo giao thông

**V. Kết luận**

Chỉ với 50 epochs thì kết quả thực nghiệm có thể thấy rằng mô hình không hoàn toàn chính xác khi thực hiện dự đoán 58 biển báo, sẽ có một số biển báo nhận diện sai so với thực tế. Kết quả của quá trình huấn luyện mô hình YOLOv8 với 50 epochs cho thấy sự tiến bộ đáng kể trong nhiệm vụ phát hiện và định vị vật thể. Mặc dù một số chỉ số như mAP50-95 và mAP50 có vẻ không quá cao (lần lượt là 68% và 82%), nhưng chúng vẫn đủ cao để cho thấy khả năng tổng quát hóa của mô hình là vẫn ổn. Điều này cho thấy mô hình YOLOv8 có khả năng phát hiện vật thể ở nhiều kích thước và hình dạng khác nhau. Các chỉ số precision và recall cũng đều ổn định, với giá trị lần lượt là 75% và 82%. Điều này cho thấy mô hình có khả năng nhận diện và định vị vật thể một cách chính xác và đáng tin cậy. Tóm lại, kết quả đạt được là khá tốt, vẫn cần có thời gian để cải thiện. Và để cải tiến hiệu suất của mô hình thì có các phương pháp cải tiến như tăng cường các dữ liệu cần dùng cho huấn luyện, điều chỉnh các tham số như nâng số epochs lên 100 hoặc cao hơn, và thử nghiệm các kiến trúc mô hình mới. Quá trình này có thể giúp cải thiện hiệu suất của mô hình và đảm bảo rằng nó có khả năng phát hiện, nhận diện và định vị vật thể tốt nhất có thể trong nhiều điều kiện, tình huống khác, hình dạng khác nhau, và điều này sẽ hỗ trợ cho dự án xu tự hành có thể xử lý một cách tốt hơn trong quá trình di chuyển xe trong môi trường thực tế.

**VI.Tài liệu tham khảo**

1. 09 Mar 2020 - phamdinhkhanh - YOLO You Only Look Once: [Khoa học dữ liệu (phamdinhkhanh.github.io)](https://phamdinhkhanh.github.io/2020/03/09/DarknetAlgorithm.html)
2. .IEEE.pdf. (n.d.-a). https://www.iitrpr.ac.in/library/pdf/IEEE.pdf
3. Duy, L. V. (2023, July 4). *Street traffic signs in Vietnam-YOLOV8*. Kaggle: https://www.kaggle.com/datasets/lvnduy/street-traffic-signs-in-vietnam-yolov8
4. Roboflow. (n.d.). *Notebooks/notebooks/train-yolov8-object-detection-on-custom-dataset.ipynb at main · roboflow/notebooks*. GitHub.: <https://github.com/roboflow/notebooks/blob/main/notebooks/train-yolov8-object-detection-on-custom-dataset.ipynb>
5. *Roboflow docs*. Annotation Tools. (n.d.): https://docs.roboflow.com/annotate/annotation-tools
6. Ultralytics. (2024, March 3). *Python*. Python - Ultralytics YOLOv8 Docs.: <https://docs.ultralytics.com/usage/python/>
7. Wang, X., Gao, H., Jia, Z., & Li, Z. (2023, October 10). *BL-Yolov8: An improved road defect detection model based on yolov8*. MDPI. https://www.mdpi.com/1424-8220/23/20/8361
8. Ultralytics. (2024a, March 2). *Yolo Performance Metrics*. YOLO Performance Metrics - Ultralytics YOLOv8 Docs. https://docs.ultralytics.com/guides/yolo-performance-metrics/#object-detection-metrics