

NHẬN DIỆN XE HƠI DỰA TRÊN HÌNH ẢNH

Vũ Huy Hoàng - 21522104

Đoàn Tiến Đạt - 21521933

Tóm tắt

- Lớp: CS519.P11
- Link Github của nhóm:
- Link YouTube video:



Vũ Huy Hoàng - 21522104



Đoàn Tiến Đạt - 21521933

Giới thiệu



Mỗi loại xe hơi khác nhau có những đặc trưng riêng khác nhau, từ đó cấu thành những yếu tố để phân loại

Mục tiêu

Sử dụng mô hình phát hiện đối tượng **YOLO (You Only Look Once)** để định vị và nhận diện xe trong các khung hình một cách nhanh chóng và chính xác, ngay cả trong môi trường giao thông phức tạp và đông đúc. Đồng thời, tích hợp mô hình phân loại hình ảnh **ResNet (Residual Network)** để xác định chi tiết kiểu dáng (sedan, SUV, coupe) và thương hiệu xe (Toyota, Honda, BMW,...) dựa trên các đặc điểm như kích thước, tỷ lệ, thiết kế cửa sổ, đèn pha và lưới tản nhiệt. Hệ thống được tối ưu hóa bằng cách sử dụng mô hình kết hợp (Hybrid Model) giữa YOLO và ResNet để đảm bảo hiệu quả trong cả phát hiện đối tượng và phân loại chi tiết, đặc biệt với các tập dữ liệu không đồng nhất hoặc chứa nhiễu.

Ứng dụng công nghệ học sâu (Deep Learning) và trí tuệ nhân tạo (AI), Tận dụng các mô hình học sâu tiên tiến để đảm bảo hệ thống có khả năng thích nghi với các điều kiện thực tế, nâng cao khả năng tự động hóa trong giám sát giao thông, quản lý bãi đỗ xe, và phân tích dữ liệu giao thông.

Kiểm tra hiệu suất của hệ thống với nhiều tập dữ liệu khác nhau, so sánh với các phương pháp truyền thống để đảm bảo hệ thống có thể được ứng dụng vào nhiều lĩnh vực thực tế, từ giám sát giao thông đến quản lý đô thị thông minh.

Mục tiêu tổng thể là tạo ra một hệ thống nhận diện xe hơi không chỉ chính xác và hiệu quả mà còn dễ triển khai trong các ứng dụng thực tế, góp phần nâng cao hiệu suất của các hệ thống quản lý giao thông và đô thị thông minh.

Nội dung và Phương pháp

Nội dung:

- Nghiên cứu và trình bày các phương pháp nhận diện xe hơi truyền thống và hiện đại, bao gồm sử dụng công nghệ xử lý hình ảnh cơ bản và các mô hình học sâu (Deep Learning).
- Trình bày đặc điểm và cách hoạt động của các mô hình phát hiện đối tượng **YOLO (You Only Look Once)**, mô hình phân loại hình ảnh **ResNet (Residual Network)** và mô hình kết hợp (Hybrid Model).
- Mô tả chi tiết dữ liệu hình ảnh được sử dụng, bao gồm kích thước tập dữ liệu, loại dữ liệu (giao thông đông đúc, đường vắng, ảnh chứa nhiều xe hoặc ít xe), và các kỹ thuật tiền xử lý để làm sạch và tăng cường dữ liệu (Data Augmentation).

Nội dung và Phương pháp

Nội dung:

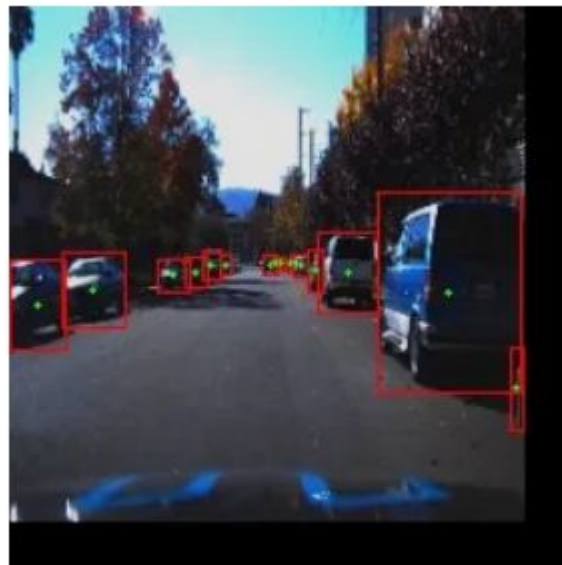
- Tích hợp các mô hình YOLO và ResNet để phát hiện và phân loại xe hơi. Phân tích cách kết hợp hai mô hình nhằm tối ưu hóa hiệu suất nhận diện và tốc độ xử lý.
- Triển khai các thử nghiệm trên tập dữ liệu thực tế để đánh giá độ chính xác, tốc độ xử lý và khả năng phát hiện xe trong các điều kiện khác nhau (ánh sáng yếu, nhiều xe, che khuất).
- Đánh giá hiệu quả của hệ thống mới so với các phương pháp nhận diện xe hơi truyền thống, về độ chính xác, thời gian xử lý, và khả năng thích nghi với các môi trường dữ liệu không đồng nhất.

Nội dung và Phương pháp

Phương pháp:

1. Xử lý và tiền xử lý dữ liệu hình ảnh

Sử dụng các kỹ thuật như cắt ảnh (cropping), điều chỉnh ánh sáng, và tăng cường dữ liệu (Data Augmentation) để chuẩn bị tập dữ liệu chất lượng cao, giúp mô hình có khả năng nhận diện tốt hơn trong các trường hợp phức tạp.



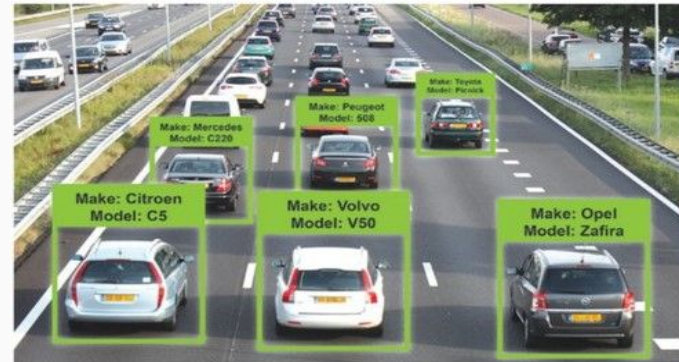
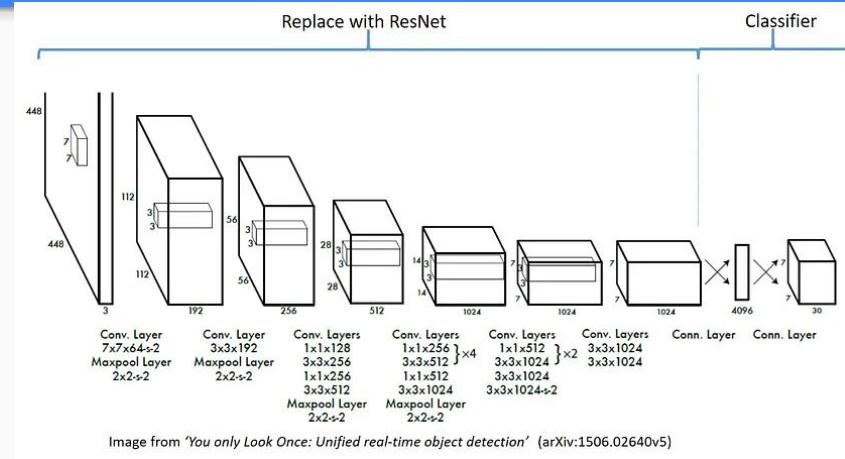
Augmented — scaled and SV adjusted

Nội dung và Phương pháp

Phương pháp:

2. Sử dụng mô hình học sâu (Deep Learning)

Áp dụng mô hình **YOLO** để phát hiện đối tượng với tốc độ nhanh và chính xác, kết hợp với **ResNet** để phân loại chi tiết các thông tin về kiểu dáng và thương hiệu xe. Các mô hình được huấn luyện với tập dữ liệu đa dạng nhằm đảm bảo hiệu quả trong các điều kiện thực tế.



Nội dung và Phương pháp

Phương pháp:

3. Đánh giá hiệu quả hệ thống bằng các chỉ số chuẩn

Sử dụng các chỉ số như **Precision**, **Recall**, **F1-Score**, và **mAP (mean Average Precision)** để đánh giá hiệu quả phát hiện và phân loại xe của hệ thống. Thực hiện phân tích so sánh hiệu suất trên nhiều tập dữ liệu khác nhau để kiểm tra khả năng mở rộng.

		Precision (Pr) $\frac{TP}{TP + FP}$	Recall (Rc) $\frac{TP}{TP + FN}$	F1-Score $\frac{2 \cdot Pr \cdot Rc}{Pr + Rc}$
A		$\frac{1}{1 + 2} = 33\%$	$\frac{1}{1 + 1} = 50\%$	$\frac{2 \cdot 33\% \cdot 50\%}{33\% + 50\%} = 40\%$
B		$\frac{2}{2 + 0} = 100\%$	$\frac{2}{2 + 0} = 100\%$	$\frac{2 \cdot 100\% \cdot 100\%}{100\% + 100\%} = 100\%$
C		$\frac{0}{0 + 2} = 0\%$	$\frac{0}{0 + 2} = 0\%$	$\frac{2 \cdot 0\% \cdot 0\%}{0\% + 0\%} = 0\%$
D		$\frac{2}{2 + 2} = 50\%$	$\frac{2}{2 + 0} = 100\%$	$\frac{2 \cdot 50\% \cdot 100\%}{50\% + 100\%} = 67\%$
E		$\frac{1}{1 + 0} = 100\%$	$\frac{1}{1 + 1} = 50\%$	$\frac{2 \cdot 100\% \cdot 50\%}{100\% + 50\%} = 67\%$

Kết quả dự kiến

- Hệ thống dự kiến đạt được độ chính xác cao trong việc phát hiện và phân loại xe hơi từ hình ảnh, đặc biệt trong các môi trường phức tạp như giao thông đông đúc, ánh sáng yếu, hoặc ảnh có nhiều xe che khuất. Mô hình YOLO được kỳ vọng mang lại tốc độ xử lý nhanh, trong khi ResNet đảm bảo độ chính xác trong phân loại chi tiết.
- Với hiệu suất tốt, hệ thống có thể được áp dụng vào các ứng dụng thực tế như giám sát giao thông, quản lý bãi đỗ xe, hoặc phân tích dữ liệu giao thông. Điều này giúp cải thiện hiệu quả vận hành và giảm tải công việc cho con người.
- Hệ thống có thể dễ dàng mở rộng để nhận diện thêm các thông tin khác như biển số xe, màu sắc, hoặc trạng thái (xe đang di chuyển hay dừng). Điều này đảm bảo tính linh hoạt và tiềm năng phát triển trong các nghiên cứu và ứng dụng tiếp theo.
- Phân tích kết quả cũng dự kiến làm rõ các hạn chế của hệ thống, chẳng hạn như hiệu suất giảm trong trường hợp dữ liệu bị nhiễu hoặc thiếu thông tin. Từ đó, bài nghiên cứu sẽ đề xuất các hướng cải tiến như tăng cường dữ liệu đầu vào hoặc sử dụng mô hình kết hợp phức tạp hơn.

Tài liệu tham khảo

- [1] **Chih-Hao Chu, Jian-Jiun Ding, Yuan Kang Lee**: "Improved YOLO Algorithm Based on Intensity Distribution Analysis for Nighttime Vehicle Recognition in Driving Recorder." In *Proceedings of the 8th International Conference on Graphics and Signal Processing (ICGSP 2024)*, Tokyo, Japan, June 14-16, 2024, pages 1-8. ACM, 2024. ISBN 979-8-4007-1702-4.
- [2] **Georgios Tzedakis, Eleftheria Tzamali, Emmanouil G. Spanakis, Marios Antonakakis, Michalis E. Zervakis, Vangelis Sakkalis**: "Comparing YOLO-based Detectors for Pedestrian and Car Detection in Aerial Static Video: An Evaluation of Generalization Capacity and Performance." In *Proceedings of the 2023 International Conference on Imaging Systems and Techniques (IST 2023)*, Copenhagen, Denmark, pages 1-5. IEEE, 2023.
- [3] **Zhongjie Tang, Chunhua Pan, Xuegang Zhang**: "A Thangka Classification Study Based on Omni-Dimensional Dynamic Convolution and ResNet." In *Proceedings of the 8th International Conference on Graphics and Signal Processing (ICGSP 2024)*, Tokyo, Japan, June 14-16, 2024, pages 18-23. ACM, 2024. ISBN 979-8-4007-1702-4.
- [4] **Zhan Wang, LiYang Wang, Yan Wang, Shuang Xia, Yue Fan, ZongWei Huang**: "An Object Extraction Algorithm Based on Ellipse Detection." In *Proceedings of the 8th International Conference on Graphics and Signal Processing (ICGSP 2024)*, Tokyo, Japan, June 14-16, 2024, pages 14-17. ACM, 2024. ISBN 979-8-4007-1702-4.